

# Quick Job Entry or Long-Term Human Capital Development? The Dynamic Effects of Alternative Training Schemes\*

ADERONKE OSIKOMINU

(University of Zurich, Albert-Ludwigs-University Freiburg, CESifo and IZA)

This draft: May 10, 2012

**Abstract:** This paper investigates how precisely short-term, job-search oriented training programs as opposed to long-term, human capital intensive training programs work. We evaluate and compare their effects on time until job entry, stability of employment, and earnings. Further, we examine the heterogeneity of treatment effects according to the timing of training during unemployment as well as across different subgroups of participants. We find that participating in short-term training reduces the remaining time in unemployment and moderately increases job stability. Long-term training programs initially prolong the remaining time in unemployment, but once the scheduled program end is reached participants exit to employment at a much faster rate than without training. In addition, they benefit from substantially more stable employment spells and higher earnings. Overall, long-term training programs are well effective in supporting the occupational advancement of very heterogeneous groups of participants, including those with generally weak labor market prospects. However, from a fiscal perspective only the low-cost short-term training schemes are cost efficient in the short run.

**Keywords:** training, program evaluation, duration analysis, dynamic treatment effects, multiple treatments, active labor market policy

**JEL:** J 64, C 41, J 68, I 28

Correspondence: Department of Economics, University of Zurich, Mühlebachstrasse 86, 8008 Zurich, Switzerland, Email: [aderonke.osikominu@econ.uzh.ch](mailto:aderonke.osikominu@econ.uzh.ch)

---

\*I am grateful to the editor Imran Rasul and three anonymous referees for their very helpful comments and suggestions. The paper also benefited from comments by Christian Dustmann, Bernd Fitzenberger, Volker Grossmann, Krisztina Kis-Katos, Marie Paul, Jeff Smith, Gerard van den Berg, and participants at various conferences, workshops and seminars in Amsterdam, Freiburg, Graz, London, Mannheim, Milan, Nottingham, Nuremberg, Washington and Zurich. This paper is part of the project “Employment effects of further training programs 2000-2002 – An evaluation based on register data provided by the Institute of Employment Research (IAB)” (IAB project number 6-531.1A). Special thanks to Stefan Bender for his efforts to make the data available. Financial support by the IAB and the German Excellence Initiative is gratefully acknowledged.

# 1 Introduction

Job training programs are the core of active labor market policy in most advanced countries. They aim at integrating the jobless and economically disadvantaged into the labor market. Countries like the US and Germany have been operating training policies for many decades. Yet the success of different approaches – in particular, human capital development as opposed to work-first strategies – remains a controversial issue. Long-term, human capital intensive training schemes provide comprehensive instruction in occupational skills. While they focus on improvement in the productivity of the unemployed, they usually do not aim at rapid reemployment. Moreover, they are relatively expensive. In contrast, job search assistance programs comprise job readiness training and instruction in job search skills. They are based on the idea that occupational skills are best acquired on the job. Consequently, they focus on quick job entry. A major advantage is their low cost. This approach has been emphasized in policy reforms implemented in 1996 in the US and the late 1990s to early 2000s in other advanced countries.<sup>1</sup> However, the limited set of skills provided may not be sufficient to address structural mismatch and to improve employment stability in the long run, a compelling concern in view of sweeping shifts of labor demand that disadvantage less-skilled workers.<sup>2</sup> After the sharp downturn of the 2008/2009 recession, many economists call for an expansion of public sponsored training programs to combat increasing rates of long-term joblessness among less-skilled persons.<sup>3</sup>

What is the best design of job training for the unemployed and the economically disadvantaged? This paper sets out to investigate how precisely short-term, job-search oriented training and long-term, human capital intensive training programs work. We examine and compare their effects on time until job entry, stability of employment, and earnings within a unified framework. For this purpose, we model the full path of transitions between different labor market states and training programs over time. Our dynamic approach has several important advantages over conventional static research designs. First, we are able to net out differential changes in the composition of treatment and comparison group persons over time. This is necessary in order to separately assess the impacts of training on unemployment and employ-

---

<sup>1</sup>See e.g. Blank (2002), European Commission (2002), OECD (2005, ch. 4; 2006) and section B.2 in the Online Appendix.

<sup>2</sup>See e.g. Acemoglu and Autor (2011), Autor et al. (2008), Dustmann et al. (2009) and Spitz-Oener (2006) for evidence on long-term shifts in the structure of wages and employment especially in the US and Germany.

<sup>3</sup>See e.g. Elsby et al. (2010) and the accompanying comments as well as the article “America’s Jobless Men – Decline of the Working Man” in *The Economist*, April 28, 2011.

ment spells and to obtain impact estimates on employment stability and wages that are not biased by systematic differences between treated and comparison persons who take up a job. Second, we are able to exactly align treated and comparison persons with respect to their prior unemployment experience and calendar time.<sup>4</sup> Thus, we avoid that e.g. differences in labor market conditions upon reemployment confound employment impacts of training. Third, our continuous-time framework avoids specification issues that arise as a consequence of aggregating along the time dimension.

A dynamic approach is also necessary because in many countries with comprehensive active labor market policies, like in our case of Germany, program assignment is not a static, one-time event. Rather, it is a recurring decision that is dynamically related to the success of job search. Job-seekers with longer unemployment spells are more likely to end up receiving treatment because they have been exposed to the assignment process for a longer period of time.<sup>5</sup> We therefore explicitly specify the accumulation of information over time and model the dynamic process of treatment assignment jointly with the transitions between labor market states. This strategy allows us, in addition, to analyze of the effect of the timing of program participation during unemployment – an important dimension of effect heterogeneity that has been neglected in the literature so far.

Further, we investigate the heterogeneity of treatment effects not only according to the timing of training during unemployment but also across different subgroups of participants. In particular, we examine how different occupational groups fare with the two training programs. This allows us to address the question whether job training programs are an effective tool to mitigate the negative effects of structural labor demand shifts. We give a detailed picture of the dynamics of training impacts, provide summary statistics of overall impacts, e.g. in terms of mean and median outcomes, and conduct a cost-benefit assessment. We thus provide an exceptionally detailed account of these two widely used active labor market programs. Such evidence is important in order to optimally allocate existing training programs and to devise new, improved schemes.

---

<sup>4</sup>Ashenfelter (1978), Heckman, Ichimura, and Todd (1997), and Heckman and Smith (1999), among others, stress the importance to carefully align treated and comparison persons with respect to their pre-treatment outcomes.

<sup>5</sup>See Fredriksson and Johansson (2008) for a formal analysis of dynamic program starts. A similar situation arises if there is a delay between the moment of (random) assignment and program start. Some persons may drop out of the treatment group before the start of the intervention because they already found a job on their own, see Heckman et al. (1999) for evidence of substantial dropout rates in major US social experiments. Abbring and van den Berg (2005) formally analyze the value of randomization at some baseline point if actual enrolment in treatment is dispersed over time.

To provide the conceptual backdrop for our empirical analysis, we introduce a novel dynamic potential outcome framework based on the theory of continuous-time counting processes. We formally describe the dynamic relationship between treatment times, potential and actual outcome times based on the notions of past, present, and future. We formulate the assumptions needed to solve the evaluation problem in a dynamic setup in which the conditioning events are sequentially updated over time. We then develop our dynamic causal model as a semi-Markov process for the transitions between the different labor market states and training programs. In this way, we are able to extend the timing-of-events approach by Abbring and van den Berg (2003a) to a setting that involves sequential updating of information, multiple treatments (short-term and long-term training), multiple outcomes (unemployment duration, employment duration, and earnings), and treatment effects that depend in a general way on the past unemployment experience, the treatment time, observed covariates and unobserved heterogeneity.

Our identification strategy relies on the conditional no-anticipation and independence assumptions as well as on results from the literature on mixed hazard rate models. Under the conditional no-anticipation assumption, the current unemployment outlook is the same for any two potential outcomes associated with different future treatment times conditional on time-constant unobservables, the elapsed unemployment experience, and the covariate history up to the current time. Under the conditional independence assumption, the current unemployment outlook of potential outcomes associated with different treatment times are conditionally independent of treatment decisions taken up to the current time. Taken together the two assumptions can be interpreted as conditional sequential randomization into treatment. The combination of three factors makes them a credible basis for causal analysis in our case. First, we focus on a homogenous sample of job-seekers with a continuous prior employment record. Second, we exploit that the allocation of training programs is driven by the short-term supply of training slots as well as private information of the caseworker. Third, we adopt a flexible modeling strategy that relies on large sample sizes, repeated observations per person and a rich set of time-constant as well as time-varying covariates.

Under conditional no-anticipation and independence, we can then study identification of the joint distributions needed for causal inference on training impacts. Specifically, we view the successive transitions from unemployment and employment as a sequence of competing and single risks models with lagged occurrence dependence and lagged duration dependence in the higher order transition rates. We apply results from the literature on duration models with proportionate unobserved

heterogeneity in the hazard rates, dependent competing risks and lagged duration dependence, i.e. Abbring and van den Berg (2003b), Brinch (2007), Honoré (1993), and Horny and Picchio (2010). Hence, our estimated treatment effects are identified semiparametrically. Our large and exceptionally informative data allow us to dispense with some important functional form assumptions that are typically made in the literature and to adopt a rich and flexible modeling strategy. In particular, we exploit that we observe repeated realizations per person and time-varying covariates.<sup>6</sup> Thus, we are able to overcome important limitations of existing evaluation studies using duration methods for program evaluation.

The empirical analysis in this paper uses unique administrative data for Germany. Germany is an interesting case to study because its recent reforms and developments in the field of labor market policy closely follow the recommendations formulated in the international policy debate in the mid-1990s in view of high unemployment levels especially in the European countries (cf. European Commission, 2002, on the “European Employment Strategy” and OECD, 2006, on the “OECD Jobs Strategy”). Average public spending on active labor market policies per unemployed relative to per capita GDP in the years 2000 and 2001 is around 33% in Germany which is at the mean across OECD countries (OECD, 2006, ch. 7).<sup>7</sup> Traditionally, long-term, human capital development oriented training programs dominated active labor market policy in Germany. Since the turn of the millennium, short-term, job-search oriented training programs have gained in importance.<sup>8</sup>

Our main findings are as follows. Participating in short-term training reduces the remaining time in unemployment and has moderate positive effects on subsequent job stability. Long-term training programs initially prolong the remaining time in unemployment. However, once the scheduled program end is reached participants exit to employment at a much faster rate than without training. Moreover, they benefit from substantially more stable employment spells and higher earnings. Further, our findings point to the possibility of improving the efficiency of long-term training programs through a careful targeting. Specifically, we find that the opportunity cost of participating in terms of prolonged unemployment is lower for people with lower chances to exit unemployment on their own, e.g. long-term unemployed and low-skilled. Persons without formal education degree and persons previously working in low- and medium-skilled manual occupations reap particu-

---

<sup>6</sup>With these data features, identification only relies on separability of the hazard rate in the unobserved heterogeneity term.

<sup>7</sup>The corresponding figure is 7% for the US, and 51% for Sweden.

<sup>8</sup>See section B.2 in the Online Appendix.

larly high gains with respect to earnings. Persons previously working in medium-skilled analytic and interactive jobs show substantial gains in terms of employment stability. Overall, long-term training programs are well effective in supporting the occupational advancement of very heterogeneous groups of participants, including those with generally weak labor market prospects. However, from a fiscal perspective only the low-cost short-term training schemes are cost efficient in the short run. In light of the recent policy shifts favoring short, job-search oriented training programs a more balanced use of both types of training seems warranted. In fact, while the shorter programs may appeal for their cost efficiency the longer programs seem more effective in tackling structural deficits.

The remainder of the paper is organized as follows. The next section discusses related literature in the field. Section 3 lays out our evaluation approach and section 4 describes the implementation. Section 5 presents our causal estimates of training impacts, and section 6 concludes. Sections A through I in the Online Appendix contain further information on the identification analysis, the institutional context, the data source, descriptive evidence, the variables used in the estimation and the estimated models.

## 2 Related Literature

The literature on microeconomic evaluations of job training programs has been growing rapidly since the late 1990s as the political demand for rigorous scientific evaluations of labor market programs increased also in the continental European countries. Evidence reviewed in various surveys and meta-analyses e.g. by Heckman et al. (1999), Greenberg et al. (2003), Kluve (2010), and Card et al. (2010) suggests that human capital intensive training programs show modestly positive effects in the medium run, with the percentage gain in employment generally exceeding that in earnings. Job search assistance programs yield favorable employment effects in the short run. In the following, we highlight evaluation studies that are most closely related to ours.

We start with studies applying experimental or matching methods. The vast majority of evaluation studies analyzes short- to medium-term outcomes of job training programs, which may bias findings towards a more favorable assessment of shorter programs, such as job search training, compared to longer ones. A few studies investigate long-term impacts, among them Couch (1992) and Hotz et al. (2006) for the US as well as Fitzenberger et al. (2008) and Lechner et al. (2011) for Germany.

Their results suggest that, while positive effects of human capital intensive training appear only with a delay, they persist over periods as long as eight or nine years. Hotz et al. (2006) and Dyke et al. (2006) study the differential effects of training programs aiming at human capital development and job search assistance, respectively, in the US. They find that job search assistance programs improve employment prospects in the short run whereas more intensive training programs initially lead to employment and earnings losses. Over the longer term, human capital intensive training tends to be more effective than job search assistance.

While experimental and matching estimators, that closely mimic the experimental design, are appealing for their straightforward interpretation, they also have some limitations. As argued above, the experimental setup is inherently static and does not take into account the dynamics of program assignment and of labor market outcomes.<sup>9</sup> In order to better study the labor market dynamics associated with program participation some researchers have used event history models to evaluate training programs. The earlier duration literature – notably Gritz (1993), Ham and Lalonde (1996) and Eberwein et al. (1997) for the US and Bonnal et al. (1997) for France – focuses on modeling the dynamic selection into different labor market states. While in Ham and Lalonde (1996) training is randomly assigned to a stock sample of eligible persons, Gritz (1993), Bonnal et al. (1997), and Eberwein et al. (1997) jointly model transitions into program states and outcome states. This early duration literature conceives program participation as a separate labor market state and models treatment effects with an indicator for past program participation in subsequent transition rates. These studies find mixed effects on subsequent unemployment spells and mostly positive impacts on subsequent employment spells.

More recently, Abbring and van den Berg (2003a) develop an econometric framework to jointly analyze dynamically assigned programs and the probability of survival in a baseline state. They prove identification of dynamic treatment effects that depend on observed covariates and either the elapsed time since treatment or unobserved heterogeneity.<sup>10</sup> Compared to the early duration literature the more recent timing-

---

<sup>9</sup>There exist a number of intermediate strategies between a fully dynamic approach and a completely static design. Matching estimators have been adapted to account for dynamic program assignment, see e.g. Sianesi (2004, 2008), Dyke et al. (2006), and Fitzenberger et al. (2008). Other approaches combine matching with dynamic outcomes such as transition rates and survival times, see e.g. Bergemann et al. (2009) and de Luna and Johansson (2009). See Abbring and Heckman (2007), Abbring and van den Berg (2004), and Crépon et al. (2009) for a methodological overview and comparison of different evaluation approaches.

<sup>10</sup>Richardson and van den Berg (2006) extend the timing-of-events approach proving identification of dynamic treatment effects that are a proportional function of time since treatment, observed covariates and unobserved heterogeneity. Abbring (2007) develops a framework based on mixed proportional hazard rate models that involves multiple baseline states and competing destination

of-events approach explicitly models the dynamics of program starts. Assignment to program and exit from the baseline state are viewed as two competing risks. If exit from the baseline state occurs first, the waiting time until treatment is censored. Conversely, if assignment to treatment occurs first, the exit rate from the baseline state may change as a consequence of the treatment. In addition, the timing-of-events approach highlights the importance of no-anticipation of future treatments for identification. Our dynamic causal model is similar to that of Abbring and van den Berg (2003a), but we use a counting process framework to describe the current evolution of potential outcome processes with respect to past and future treatment events. We adapt the conditional no-anticipation and independence assumptions to a setting in which the conditioning events are sequentially updated over time. Further, we consider multiple types of treatments and multiple outcomes as well as treatment effects that depend in a general way on the past unemployment experience, the treatment time, observed covariates and unobserved heterogeneity.

A small literature applies the timing-of-events approach to evaluate occupational training (Crépon et al., 2011, for France, Lalive et al., 2008, for Switzerland, Hujer et al., 2006b, for Germany, and Richardson and van den Berg, 2006, for Sweden) or active job-search programs (Crépon et al., 2005, for France, Hujer et al., 2006a, for Germany, and Weber and Hofer, 2004a, b, for Austria). The vast majority focuses on impacts on the duration in the initial unemployment spell, only Crépon et al. (2005) and Crépon et al. (2011) also study unemployment recurrence. The evidence in these studies suggest that job search training reduces unemployment duration, while more intensive training programs tend to increase it. Crépon et al. (2005) and Crépon et al. (2011) find that both program types have a beneficial effect on unemployment recurrence.

A major advantage of our study is its exceptionally large and informative administrative data. Thus, we can relax important functional form assumptions that are typically made in the duration literature. We adopt a flexible modeling strategy and provide a detailed account of job-search oriented and human capital intensive training with respect to the outcomes unemployment duration, employment duration and earnings. As a further contribution, we examine the heterogeneity of treatment effects across a range of important observed characteristics such as education and occupation.

---

states.



## 3 Evaluation Approach

### 3.1 Dynamic Potential Outcome Framework

Consider the following setup. People can either be unemployed or employed. While unemployed, they may be assigned to a training program. In particular, the job-seeker and the caseworker at the local employment agency meet repeatedly during the unemployment spell. At any such occasion, the caseworker decides whether to assign the job-seeker to a program or to postpone participation to the future, waiting further how job search evolves. Somebody who has not participated, say, until day 80 of his unemployment spell may still enrol later. If, however, he starts a new job at day 81 he would not be eligible for participation anymore. Thus, a job-seeker in open, i.e. untreated, unemployment is exposed to two risks that compete to end open unemployment: start of a training program and start of a new job. How can we evaluate the effect of participating in training on the duration of unemployment in this setting?

We adopt a continuous-time version of the potential outcome approach to program evaluation (Neyman, 1923, Roy, 1951, Rubin, 1974).<sup>11</sup> The different starting times of training during unemployment represent a continuum of mutually exclusive treatments  $s \in \mathbb{R}_+ \cup \{\infty\}$ .  $s = \infty$  corresponds to the case that training never starts. We denote by  $T^*(s) \in \mathbb{R}_+$  the potential duration in unemployment until exit to employment that prevails if training starts at time  $s$  of unemployment.<sup>12</sup> Let  $T \in \mathbb{R}_+$  be the actual unemployment duration. For training starting at  $s$ , the actual outcome time is identical to the potential outcome time associated with  $s$ , i.e.  $T \equiv T^*(s)$ , whereas the other potential outcome times  $T^*(s')$ ,  $s' \neq s$ , are counterfactual.

We express the random time spent in unemployment as a continuous-time stochastic process. This will allow us to formally describe how past and future treatment times affect the current unemployment outlook conditional on the information accumulated until the current period. Let  $u$  denote the elapsed time since the beginning of unemployment. The actual unemployment duration  $T$  generates a counting process  $\mathbf{Y} \equiv \{Y(u), u \in \mathbb{R}_+\}$  such that  $Y(u) = 0$  if  $u \leq T$  and  $Y(u) = 1$  if  $u > T$ . Analogously, we denote the counting process associated with the potential unemployment duration  $T^*(s)$  as  $\mathbf{Y}^*(s) \equiv \{Y^*(s, u), u \in \mathbb{R}_+\}$  such that

---

<sup>11</sup>See Abbring and van den Berg (2003a) and Lok (2008) for related approaches in continuous time. Abbring and Heckman (2007) provide an overview of different dynamic evaluation approaches.

<sup>12</sup>We assume that all potential outcomes  $\{T^*(s), 0 \leq s \leq \infty\}$  are absolutely continuous random variables.

$Y^*(s, u) = 0$  if  $u \leq T^*(s)$  and  $Y^*(s, u) = 1$  if  $u > T^*(s)$ .<sup>13</sup> Further, let  $dY(u)$  be the differential of  $Y(u)$  in a small interval  $[u, u + du)$ . Let  $u^-$  denote the time just before  $u$ . The probability that an exit to employment occurs in the small interval  $[u, u + du)$  conditional on survival in unemployment until just before  $u$  is given by  $\Pr(dY(u) = 1 | Y(u^-) = 0) = \Pr(T \in [u, u + du) | T \geq u) \equiv \theta(u)du$ , where  $\theta(u)$  denotes the hazard rate or intensity process. In order to model this probability conditional on the information available up to the current period, we introduce the filtration  $\mathcal{I}(u^-)$  that collects the information accumulated until just before  $u$ . Let  $\{x(u), u \in \mathbb{R}_+\}$  be a vector-valued stochastic process of time-varying random variables and  $v$  a vector of time-constant random variables.<sup>14</sup> While  $x(u)$  represents covariates that can be observed by both the job-seeker and the researcher,  $v$  is not observable for the researcher. Now  $\mathcal{I}(u^-)$  is defined as  $\sigma\{Y(r), x(r), v, 0 \leq r < u\}$ , where  $\sigma\{\cdot\}$  is the  $\sigma$ -algebra generated by the events  $\{Y(r), x(r), v, 0 \leq r < u\}$ .  $\mathbf{Y}$  has an intensity process with respect to the filtration  $\mathcal{I}(u^-)$  if  $\theta(u|\mathcal{I}(u^-))$  is predictable given  $\mathcal{I}(u^-)$  and, thus, the conditional expectation of  $dY(u)$  equals the conditional intensity process times the increment of  $u$ :  $E[dY(u) - \theta(u|\mathcal{I}(u^-))du | \mathcal{I}(u^-)] = \Pr[dY(u) = 1 | \mathcal{I}(u^-)] - \theta(u|\mathcal{I}(u^-))du = 0$ .

Our goal is to contrast the marginal distributions of potential outcome times associated with different treatment times. In particular, we want to evaluate the effect of starting training at  $s$  instead of  $s'$  for those who enrol at  $s$ . The potential outcome distribution associated with starting training at  $s'$  is counterfactual for those who enrol at  $s$ . In order to solve the evaluation problem we rely on two assumptions. As only unemployed persons are eligible for training the treatment times are censored for people who find a job before starting training. Therefore, we assume that conditional on the information accumulated up to just before  $u$  the probability of an exit to employment in a small interval  $[u, u + du)$  is the same across potential outcomes associated with different treatment times that lie after  $u$ . Formally, for any two potential outcome processes  $\mathbf{Y}^*(s)$  and  $\mathbf{Y}^*(s')$  that have intensity processes with respect to the filtration  $\mathcal{I}(u^-) \equiv \sigma\{Y(r), x(r), v, 0 \leq r < u\}$ , we suppose that:

(1) *(Conditional No-Anticipation)*

$$\Pr[dY^*(s, u) = 1 | \mathcal{I}(u^-)] = \Pr[dY^*(s', u) = 1 | \mathcal{I}(u^-)] \quad \forall u \leq \min(s, s').$$

Under the conditional no-anticipation assumption (1), the current probabilities of survival in unemployment for any two potential outcome processes associated with different future treatment times coincide. Hence, we can study the distribution

---

<sup>13</sup>We assume that all counting processes are right-continuous and have limits on the left.

<sup>14</sup>We suppose that the process  $\{x(u), 0 \leq u\}$  is left continuous and that  $v$  is bounded.

of the observed unemployment process given that training has not yet started to recover the potential unemployment process without training ( $s = \infty$ ).<sup>15</sup>

Second, we assume that conditional on the history of information available just before  $u$  the potential outcome processes are independent of the treatment history up to  $u$ . Formally, the random waiting time in unemployment until the start of training induces a counting process  $\mathbf{D} \equiv \{D(u), u \in \mathbb{R}_+ \cup \infty\}$  such that  $D(u) = 0$  if  $u$  is smaller or equal to the treatment time and  $D(u) = 1$  else. Denote the  $\sigma$ -algebra generated by the treatment process as  $\sigma\{D(r), 0 \leq r < u\}$ . We suppose that the potential outcome process  $Y^*(s)$  has an  $\mathcal{I}(u^-)$ -predictable intensity process that is also the intensity process of  $Y^*(s)$  with respect to the extended filtration  $\mathcal{J}(u^-) \equiv \mathcal{I}(u^-) \cup \sigma\{D(r), 0 \leq r < u\}$ , i.e.:

$$(2) \quad (\textit{Conditional Independence}) \\ \Pr[dY^*(s, u) = 1 | \mathcal{J}(u^-)] = \Pr[dY^*(s, u) = 1 | \mathcal{I}(u^-)] \quad \forall s.$$

While assumption (1) characterizes the evolution of current potential outcomes associated with future treatment times, assumption (2) refers to the relationship between current potential outcomes and past treatment events. Taken together the two assumptions can be interpreted as sequential randomization into treatment conditional on the information available up to the time of randomization. They are comparable to sequential conditional unconfoundedness assumptions invoked in sequential matching approaches (e.g. Lok, 2008, Robins, 1998, Sianesi, 2004), except that we allow the conditioning set to include time-constant unobservables. Importantly, assumptions (1) and (2) do not rule out that job-seekers can in general predict the probability of receiving training at particular points in time. Moreover, their predictions may be based on information acquired during the course of unemployment. Assumptions (1) and (2) just require that the researcher can control for all events that jointly predict treatment and outcome times. In addition, after conditioning on all relevant information, treatment assignment still needs to be a stochastic process.<sup>16</sup> In section 3.3, we substantiate the empirical support of assumptions (1) and (2) in our application.

---

<sup>15</sup>Without the conditional no-anticipation assumption, we could only identify the average of pre-treatment potential outcomes with respect to the distribution of future starting dates.

<sup>16</sup>Abbring and van den Berg (2003a) also rely on versions of the conditional no-anticipation and independence assumptions. Translated into our notation, their no-anticipation assumption states that  $\Pr[dY^*(s, u) = 1 | \mathcal{H}(u^-)] = \Pr[dY^*(s', u) = 1 | \mathcal{H}(u^-)] \quad \forall u \leq \min(s, s')$ , where  $\mathcal{H}(u^-) \equiv \sigma\{Y(r), 0 \leq r < u, x(0), v\}$ . Thus, they abstract from the possibility that the current transition probabilities of potential outcomes associated with different future treatment times depend differentially on events occurring between the start of unemployment and the current period. Translated into our notation, their conditional independence assumption states that the potential outcome processes,  $\{Y^*(s), s \in \mathbb{R}_+ \cup \infty\}$ , are independent of the treatment process,  $\mathbf{D}$ , conditional

By assumptions (1) and (2) we can link the potential probability of an exit to employment at time  $u$  to the actually observed probability:

$$\Pr[dY^*(s, u) = 1 | \mathcal{I}(u^-)] = \Pr[dY^*(s, u) = 1 | \mathcal{J}(u^-)] = \Pr[dY(u) = 1 | \mathcal{J}(u^-)].$$

Thus, we can conduct inference on the potential unemployment processes associated with different treatment times by studying the properties of the actually observed unemployment processes conditional on the respective treatment times. Further, we can express the actual outcome process and the treatment process as three underlying processes that are independent conditional on  $\mathcal{J}(u^-)$ . The first component process tracks the transition from open unemployment to employment, the second the transition from open unemployment to treatment and the third the transition from treated unemployment to employment. Thus, the waiting times in open unemployment until the start of training and until exit to employment are two competing risks. Identification of the joint distribution of the waiting times is nontrivial because we allow all three to depend on the unobserved heterogeneity  $v$ . We rely on results from the literature on mixed hazard rate and competing risks models. We describe the full model, that includes two types of training programs, and its identification in section 3.2.

Further, we want to study the impact of training on the stability of subsequent employment. We extend our framework as follows. Let  $K(u)$  be the state occupied at time  $u$  since inflow into unemployment.  $K(u) = U$  if unemployed and  $K(u) = E$  if employed at time  $u$ . The transition from origin state  $k$  to destination state  $l$ , can be modeled as a continuous-time semi-Markov process. Analogously to above, let  $T_{kl} \in \mathbb{R}_+$  denote the random time spent in state  $k$  until exit to state  $l$  and  $\mathbf{Y}_{kl}$  the corresponding counting process that tracks the transition from state  $k$  to state  $l$ . Similarly,  $\mathbf{Y}_{kl}^*(s)$  denotes the potential outcome process for the transition from state  $k$  to  $l$  prevailing if training starts at time  $s$  of unemployment. The probability that a transition from  $k$  to  $l$  occurs in a small interval  $[u, u + du)$  is given by  $\Pr[dY_{kl}(u) = 1 | Y_{kl}(u^-) = 0, K(u^-) = k] = \Pr[K(u + du) = l | T_{kl} \geq t, K(u^-) = k]$ , where  $t$  is the elapsed duration in the current state  $k$ .

With some small modifications, assumptions (1) and (2) carry over to the extended setup. In both equations, substitute  $\mathbf{Y}_{kl}^*$  for  $\mathbf{Y}^*$ . The filtrations are now defined as  $\mathcal{I}(u^-) \equiv \sigma\{Y_{kl}(r), x(r), v; 0 \leq r < u; k, l \in \{E, U\}, k \neq l\}$  and  $\mathcal{J}(u^-) \equiv \sigma\{Y_{kl}(r), D(\min(r, T_{UE})), x(r), v; 0 \leq r < u; k, l \in \{E, U\}, k \neq l\}$ . Further, we assume that the conditional no-anticipation assumption (1) holds for all

---

on the information available at time zero,  $\mathcal{I}(0) \equiv \sigma\{x(0), v\}$ . This assumption abstracts from the possibility that events occurring during the course of unemployment may jointly influence potential outcome times and treatment times.

$u \leq \min(s, s')$  if  $K(u) = U$  and for all  $s, s' > T_{UE}$  if  $K(u) = E$ . Thus, the survival experience is the same for potential employment processes associated with different treatment times that have not been realized during the preceding unemployment spell. Under assumptions (1) and (2), we can recover the potential employment processes associated with different treatment times from the actual employment process conditional on the respective treatment times. Hence, we view our dynamic causal model as a semi-Markov process that involves a competing risks model for the waiting times in open unemployment until exit to employment and training and single risk models for the transitions from treated unemployment to employment as well as from treated and nontreated employment to unemployment. In addition, we consider a second model with earnings instead of employment duration. In this case, we model the conditional distribution of earnings as a hazard rate together with the hazard rates from unemployment into training and into employment.<sup>17</sup> We describe our setup and the identification in detail in the next section.

### 3.2 Identification with Proportionate Unobserved Heterogeneity

We now describe our empirical setup, that comprises two treatments, i.e. short-term training and long-term training, and two outcomes, i.e. the duration until reemployment and the duration of subsequent employment or the level of subsequent earnings in terms of a continuous-time semi-Markov process. We then discuss identification of our model based on results from the literature on mixed hazard rate models with proportionate unobserved heterogeneity.

Consider an inflow sample into the initial state open unemployment  $O$ . A person in the initial state may exit to any of the three states (i) unemployment treated with short-term training  $S$ , (ii) unemployment treated with long-term training  $L$ , or (iii) employment  $E$ . A person in state  $S$  may exit to the destinations  $L$  or  $E$ . A person in state  $L$  may move to state  $E$ .<sup>18</sup> We call one such sequence of transitions from unemployment to employment a cycle. In the model with employment duration as the second outcome we also model the transition from employment,  $E$ , back to open unemployment,  $O$ , and consider repeated cycles per individual. In the alternative

---

<sup>17</sup>Donald et al. (2000) also use a hazard rate model to investigate differences between the US and Canadian conditional wage distributions.

<sup>18</sup>If short-term training is the first intervention we allow for a second treatment with long-term training because short-term training is also used to assess the professional skills of an unemployed and to define a suitable reintegration plan, which may entail a later participation in a long-term training program.

model with earnings as the second outcome, we use just the first cycle for every individual.<sup>19</sup> Figure 1 illustrates the possible pathways from open unemployment to employment.

— Insert figure 1 here. —

More formally, let  $\mathcal{K} = \{O, S, L, E\}$  denote the state space. The function  $\mathcal{Z}(k)$  assigns to each origin state  $k \in \mathcal{K}$  a set of possible destination states:

$$\mathcal{Z}(k) = \begin{cases} \{S, L, E\} & \text{if } k = O \\ \{L, E\} & \text{if } k = S \\ \{E\} & \text{if } k = L \\ \{O\} & \text{if } k = E. \end{cases}$$

For  $k \in \{O, S\}$ , the number of elements in  $\mathcal{Z}(k)$  is greater than one,  $|\mathcal{Z}(k)| > 1$ . Thus, we view a cycle as a sequence of competing risks and single risk models.

According to the Markov assumption, the probability to experience a transition from current state  $k$  to destination state  $l$  only depends on the current state  $k$  occupied. In order to differentiate e.g. between employment spells following a participation in training as opposed to open unemployment (cf. figure 1), we define the augmented state space  $\mathcal{K}^a = \{O, S(O), L(O), E(O), L(S), E(S), E(S, L)\}$ , where the states in parentheses indicate the past trajectory. We denote the corresponding function of destination states from  $k \in \mathcal{K}^a$  by  $\mathcal{Z}^a(k)$ . We model dependence on the past trajectory as lagged occurrence dependence and lagged duration dependence. With this extension we can align treated and comparison persons with respect to their unemployment experience since the beginning of the cycle. In particular, when estimating training impacts on the transition rate to employment we align treated and comparison persons in their total elapsed unemployment duration, that is for the treated the prior duration in open unemployment plus the current duration in treated unemployment. Similarly, when measuring training impacts on the transition rate from employment back to unemployment we condition on the total prior unemployment duration.

Let  $T_{kl} \in \mathbb{R}_+$  denote the random time spent in state  $k$  before exiting to state  $l$ ,  $k \in \mathcal{K}^a$ ,  $l \in \mathcal{Z}^a(k)$  and  $t$  the elapsed duration in state  $k$ . For origin states with multiple destinations we do not observe the realizations of all exit times from state  $k$  but just that of the minimum. Let  $x(t)$  denote the covariate process just before

---

<sup>19</sup>Due to the reporting rules of the employment register we cannot model transitions into and out of employment at the same time as earnings, see section 4.2.

time  $t$ ,  $\bar{x}(t) \equiv \{x(r), 0 \leq r < t\}$  the covariate path on  $[0, t)$  and  $X$  its support. We can express the conditional distribution of  $T_{kl}$ ,  $k \in \mathcal{K}^a$ ,  $l \in \mathcal{Z}^a(k)$ , in terms of the hazard rate:

$$(3) \quad \begin{aligned} \theta_{kl}(t|\tau, \bar{x}(t), v_{kl}) &\equiv \lim_{dt \rightarrow 0} \frac{\Pr[t \leq T_{kl} < t + dt | T_{kl} \geq t; \tau, \bar{x}(t), v_{kl}]}{dt} \\ &= \lambda_{kl}[t, \mathbb{1}(k \neq O)\tau, x(t)] v_{kl}, \end{aligned}$$

where  $v_{kl} \in \mathbb{R}_+$  is an unobserved heterogeneity term while  $\lambda_{kl}(\cdot) : (\mathbb{R}_+ \times \mathbb{R}_+ \times X) \rightarrow \mathbb{R}_+$  denotes the part of the hazard that is a function of observed factors, i.e. elapsed time in the current state  $t$ , observed covariates  $x(t)$  and lagged duration  $\tau$  that enters the transition rate if origin state  $k \neq O$ .

Our goal is to identify the single components of the hazard rates  $\theta_{kl}(\cdot)$  and the joint distribution of the unobserved heterogeneity terms,  $G(\mathbf{v})$ ,  $\mathbf{v} = \{v_{kl} : k \in \mathcal{K}^a, l \in \mathcal{Z}^a(k)\}$ , which then yields the joint distribution of the durations  $\{T_{kl} : k \in \mathcal{K}^a, l \in \mathcal{Z}^a(k)\}$ . We proceed in a sequential way, considering first identification of the competing risks from the initial state, then of the hazard rates from the second state given the first transition and so on. In this way we identify, for a given origin state  $k$ , the joint distribution of the unobservables  $\{v_{kl} : l \in \mathcal{Z}^a(k)\}$  together with the selection process into state  $k$ . Once we have identified all possible trajectories leading from open unemployment to employment, we can trace out the joint distribution of the unobservables  $\{v_{kl} : k \in \mathcal{K}^a, l \in \mathcal{Z}^a(k)\}$  by varying the observed arguments of the hazard rates (i.e.  $t, \tau, x(t)$ ) on a nonempty open set. This yields then the joint distribution of the survival times  $\{T_{kl} : k \in \mathcal{K}^a, l \in \mathcal{Z}^a(k)\}$ .

We describe the exact procedure and the identification results used at each stage of the unemployment-employment cycle in section A in the Online Appendix. Here, we summarize the main points. In our model with employment duration as the second outcome, our identification strategy exploits that we observe about half of the sample more than once in the initial state and that the hazard rates from treated unemployment and employment include time-varying covariates.<sup>20</sup> This allows us to avoid several restrictive assumptions, e.g. proportionality of the effects of elapsed duration, lagged duration and observed covariates, that are commonly made in the literature on mixed hazard rate models (Abbring and van den Berg, 2003b, Proposition 3; Brinch, 2007). As suggested in equation (3), we only assume separability

---

<sup>20</sup>In the hazard rates from unemployment we use time-varying regressors related to the remaining entitlement period of unemployment benefits, age, and season of the year. In the hazard rates from treated unemployment, we include in addition indicators for different time intervals relative to the planned end of the program. In the employment hazards, we consider a polynomial of age (interacted with gender), indicators for season of the year and interactions between season and industry of the last job.

of the unobserved heterogeneity terms.

In our alternative model with earnings, we model the conditional distribution of earnings as a hazard rate together with the hazard rates from unemployment into employment, short-term training, and long-term training. In this model, we do not have repeated cycles per person and the hazard rates of daily earnings do not include covariates that vary with the level of earnings. Therefore, we assume that the hazard rates are separable in their components, a flexible nonnegative function of elapsed duration,  $\lambda_{kl}^0(t)$ , of lagged duration,  $\kappa_{kl}(\mathbb{1}(k \neq O)\tau)$ , of observed covariates,  $\phi_{kl}(x(t))$ , and unobserved heterogeneity,  $v_{kl}$ , i.e.  $\theta_{kl}(t|\tau, \bar{x}(t), v_{kl}) = \lambda_{kl}^0(t)\kappa_{kl}(\mathbb{1}(k \neq O)\tau)\phi_{kl}(x(t))v_{kl}$ , and that the regressors are independent of the unobserved heterogeneity terms (Abbring and van den Berg, 2003b, Proposition 2; Honoré, 1993, Theorem 3; Horny and Picchio, 2010).

### 3.3 Empirical Support of the Conditional No-Anticipation and Independence Assumptions

The conditional no-anticipation assumption (1) and the conditional independence assumption (2) are important prerequisites for identification. Analogously to a static matching analysis their plausibility depends on the richness of our data and our ability to flexibly control for potentially time-varying confounders that jointly determine outcome and treatment times. We can to some extent dispense with time-constant observed covariates through including unobserved heterogeneity. The combination of three factors makes the conditional no-anticipation assumption a credible basis for causal analysis in our case: (i) our construction of the analysis sample, (ii) the institutional setup determining training participation, and (iii) our flexible modeling strategy based on large and informative data.

First, we construct our analysis sample in such a way that included individuals have similar recent employment histories and high incentives to comply with the instructions of the caseworker. According to the regulation of public employment services in Germany, job-seekers face tight job search requirements (section B.1 in the Online Appendix). Therefore, we focus on prime-age workers who experience a transition into unemployment after a period of continuous employment. We refrain from including individuals who start looking for a job after a period out of the labor force because they might only register at the local employment agency so long as their expected utility from receiving particular services and programs exceeds their disutility associated with tight job search requirements. As a consequence more than



90% of the persons in our sample have substantial entitlements to unemployment benefits (section D.3 in the Online Appendix). These individuals are presumably highly committed to cooperate with the local employment agency, for intrinsic reasons as well as for the fact that in case of noncompliance their generous benefits would be at risk. Overall, these sample selection criteria ensure that the individuals considered in the analysis behave in a relatively predictable way and according to the guidelines set out by public employment services.

Second, with respect to the assignment of training programs, a participation in training may take place at any point in time during the unemployment spell. Job-seekers have no entitlements regarding participation (section B.1). However, a program assignment is compulsory for the job-seeker and noncompliance may entail benefit sanctions and the exclusion from further services. The allocation of training programs depends on the short-term supply of training slots as well as private information of the caseworker. On the one hand short notice periods and belated assignments (i.e. after the official starting date of a program) are used as a work test. On the other hand they allow caseworkers to achieve a high utilization of available training capacities and at the same time to avoid disincentives for job-seekers anticipating future program participation. As caseworkers face a very high caseload the assignment is not targeted.<sup>21</sup> As supporting evidence for this practice, we find considerable variation in the timing of training starts during the unemployment spell and modest differences in key observed characteristics between participants and nonparticipants (sections D.2 and D.3). Anticipatory effects with regard to future training participation are therefore unlikely to occur.

The third important ingredient for conditional no-anticipation and independence is the data and the specification of the conditional hazard rates. Our administrative database is exceptionally rich with regard to the number of observations as well as the available covariate information. Large sample sizes allow us to condition on current and lagged duration, calendar time, observed covariates and unobserved heterogeneity in a detailed and flexible way (see also section 4.3). For instance, we use step functions to model duration dependence, which allows us to nonparametrically approximate arbitrary patterns of duration dependence. We model the impacts of training on the exit rate out of unemployment with a step function indicating different time intervals relative to the planned end of the program. In addition, we flexibly condition on calendar time at the start of the spell, current season of the

---

<sup>21</sup>The average caseload in terms of registered unemployed is about 400 unemployed per caseworker (section B.1). In addition, caseworkers counsel people who register as job-seekers but are not unemployed as well as people who do not register in the end.

year, and age of the individual. A careful specification of duration and calendar time effects is important to capture changes in the job-seeker’s information set that are a function of time.

In terms of covariates, we have access to the information collected on job-seekers when they register at the local employment agency, i.e. the Supply of Applicants database. It builds the basis for the counseling activities and assignment decisions of the caseworker. It details personal characteristics, properties of the last job, and objectives of job-search. The data also include caseworker assessments of the qualifications and experience of a job-seeker and of his/her health status. Combining this with the employment register we can characterize the last employment relationship in terms of e.g. previous earnings, occupation, industry, type of position, and the reason why the job ended. In addition, we control for region of residence and local labor market conditions. We include time-varying indicators for the remaining benefit entitlement, past and current periods of sickness as well as past and current temporary suspensions of benefit payments.<sup>22</sup>

## 4 Implementation

### 4.1 Training Programs Analyzed

Training schemes are the core of active labor market policy in Germany as in most advanced countries.<sup>23</sup> In our analysis, we focus on human capital intensive long-term training on the one hand and job-search oriented short-term training on the other hand.

Long-term training schemes comprise a variety of programs ranging from advanced vocational training and refresher courses on specific professional skills and operational techniques to comprehensive retraining in a new vocational degree within the German apprenticeship system. The former typically last between six and twelve months whereas retraining takes two to three years. Training programs may take place either in classrooms, simulated workplaces, firms or a combination thereof. Typical examples of long-term training programs include training on marketing and sales strategies, computer assisted bookkeeping, operating construction machines, and specialist courses in specific legal fields.

---

<sup>22</sup>See section E in the Online Appendix for a detailed description of the included variables.

<sup>23</sup>See section B.2 in the Online Appendix for further information on the quantitative importance of training in the context of German active labor market policy.

Short-term training courses last a couple of days to twelve weeks. Similar to the long-term training schemes, they may take place in classrooms or workplace-like environments. However, due to their shorter length their contents are less occupation specific and the human capital component is limited. Typical examples of short-term training schemes include job application training, basic computer courses, language courses and short-term internships at a simulated or real workplace. Their aim is twofold. On the one hand, they provide skills that improve and facilitate job search. On the other hand, they are employed to assess a job-seeker's abilities and his/her readiness to work or to participate in a further program.

## 4.2 Data

The empirical analysis is based on an exceptionally rich administrative database, the Integrated Employment Biographies Sample (IEBS), that has recently been made available by the Research Data Center of the German Federal Employment Agency. The IEBS is a 2.2% random sample from a merged data file containing individual records out of four different administrative registers. It comprises data on employment subject to social security contributions, receipt of transfer payments during unemployment, job search, and participation in different active labor market programs. Start and end dates of the different labor market episodes are measured with daily precision.

From this data set, we extract a sample of West Germans aged 25 to 53 who experience a transition from regular, unsubsidized employment lasting three months or longer to unemployment within the period July 1999 to December 2001. Unemployment is defined as nonemployment with at least occasional contact with the employment agency. This contact may be recorded in the data either as a benefit spell, a registered job search spell, or a period of program participation.<sup>24</sup> Unemployment spells are censored at the end date of the last contact with the employment agency if in the following three months no such contact persists. Nonemployment spells without any contact with the employment office are not considered because in that case we cannot distinguish between people looking for a job and persons out of the labor force. Also, people who are not actively looking for a job or who do not register with the employment agency are not eligible for training programs. Transitions to active labor market programs other than training are treated as independent censoring. We follow the persons in the inflow sample until the end of

---

<sup>24</sup>For 12 out of 13,859 training spells, the training spell constitutes the first contact with the employment agency recorded in the data.

December 2004, and ongoing spells are censored at that date. We model the trajectory from open unemployment to employment, which may include a participation short-term training or long-term training. In addition, we consider the transition from employment back to unemployment. Individuals in the sample may have multiple unemployment and employment spells if they experience multiple transitions between unemployment and employment.

In the alternative model with earnings, we model the pathway from open unemployment to employment and the average earnings per calendar day within the first 365 days upon reemployment. Days during which a person is not employed contribute with a zero. Earnings are censored at the social security threshold and if the time between the end of unemployment and the end of our observation period is less than 365 days. In the model with earnings, we use only the first cycle for every individual because we cannot measure earnings dynamics at the same frequency as employment dynamics. Total earnings are reported only once every calendar year in employment relationships that last longer than a year. We model the conditional density of daily earnings in the same way as the transition rates into different labor market states, i.e. using a mixed proportional hazard rate specification.

Figure 1 above illustrates the different transitions of the 45,459 people in our sample. In total, there are 87,250 unemployment spells, of which 8,279 (5,580) lead to a participation in short-term training (long-term training), and there are 56,758 employment spells. About half of the individuals in the sample experience multiple unemployment spells and about 30% have more than one employment spell. In contrast, only 8% (2%) of the treated with short-term training (long-term training) participate more than once in the same program.

Table 1 provides summary statistics on the distributions of unemployment and employment durations as well as wages, by treatment status and for the entire sample. The means in table 1 refer to the expected value conditional on the truncated distribution of  $T$  until truncation point  $\bar{t} > 0$ .<sup>25</sup> It is calculated as follows:

$$(4) \quad E[T|T \leq \bar{t}] = \frac{\int_0^{\bar{t}} r dF(r)}{F(\bar{t})} = \frac{\int_0^{\bar{t}} S(r) dr - \bar{t}S(\bar{t})}{1 - S(\bar{t})},$$

where  $F(t)$  is the cumulative distribution function and  $S(t)$  the survivor function. We estimate  $S(t)$  with the Kaplan and Meier (1958) estimator. In sum, the evidence in table 1 suggests that long-term training has beneficial effects on employment

---

<sup>25</sup>We choose the truncation points for unemployment and employment duration according to the limits of our observation period and for daily earnings according to the social security threshold in 2000.

duration but at the same time seems to strongly increase unemployment duration. Short-term training, in contrast, tends to reduce unemployment duration but seems to have mixed effects on employment and earnings. However, it is important to stress that these patterns reflect a mixture of causal and selection effects. It remains to be seen to what extent they persist after taking into account the dynamic selection into treatments and into outcome states.<sup>26</sup>

— Insert table 1 here. —

### 4.3 Model Specification

We adopt a piecewise constant exponential model for the hazard rates. For the index functions, we use flexible, linear in parameters specifications to model the dependence on observed covariates  $x(t)$  and lagged duration  $\tau$ . We use piecewise constant specifications to model the dependence on elapsed duration  $t$  and time dependence of training impacts during unemployment. Specifically, in our model with employment, the hazard rate for the transition from state  $k$  to  $l$  is given by:

$$\begin{aligned}\theta_{kl}(t|\tau, x(t), v_{kl}) &= \lambda_{kl} [t, \mathbb{1}(k \neq O)\tau, x(t)] v_{kl} \\ &= \exp[h_{kl}(t, \mathbb{1}(k \neq O)\tau, x(t))\beta_{kl}] v_{kl},\end{aligned}$$

where  $h_{kl}(\cdot)$  is a function of the observed components elapsed time  $t$ , lagged duration  $\tau$  (that enters the hazard rate if state  $k \neq O$ ) and observed covariates  $x(t)$ , and  $\beta_{kl}$  is a conformable coefficient vector. In our alternative model with earnings, the hazard rate for the transition from state  $k$  to  $l$  is modeled as  $\theta_{kl}(t|\tau, x(t), v_{kl}) = \exp[b_{kl}(t)\gamma_{kl} + d_{kl}(\mathbb{1}(k \neq O)\tau)\delta_{kl} + f_{kl}(x(t))\zeta_{kl}] v_{kl}$ , with  $b(\cdot)$ ,  $d(\cdot)$ ,  $f(\cdot)$  some flexible functions and  $\gamma$ ,  $\delta$ ,  $\zeta$  conformable coefficient vectors.

We model the joint distribution of the unobservables,  $\{v_{kl} : k \in \mathcal{K}^a, l \in \mathcal{Z}^a(k)\}$ , with a discrete masspoint distribution that in principle allows to approximate any arbitrary discrete or continuous distribution (Heckman and Singer, 1984). In particular, we adopt a factor loading specification with two independent underlying factors,  $w_1$  and  $w_2$ :

$$v_{kl} = \exp(\alpha_{kl1} w_1 + \alpha_{kl2} w_2), \quad k \in \mathcal{K}^a, l \in \mathcal{Z}^a(k),$$

where  $\alpha_{kl1}$  and  $\alpha_{kl2}$  denote the factor loadings on the fundamental unobserved factors  $w_1$  and  $w_2$ , respectively, for a transition from state  $k$  to  $l$ . We normalize  $w_1$  and  $w_2$

---

<sup>26</sup>Sections C, D, and E in the Online Appendix provide further details on the administrative data base, additional descriptive evidence and a complete list of the covariates used in the estimations.

to have support  $\{-1, 1\}$  and in addition constrain one of the  $\alpha_{kl2}$  to equal zero.<sup>27</sup> The two-dimensional factor loading model imposes no restrictions on the covariance matrix of the unobserved heterogeneity terms. Let  $\mathbf{w} = (w_1, w_2)'$  and  $A$  be the matrix of factor loadings with rows  $A_{kl} = (\alpha_{kl1}, \alpha_{kl2})$ ,  $k \in \mathcal{K}^a, l \in \mathcal{Z}^a(k)$ . Then the variance-covariance matrix of the unobserved heterogeneity terms is given by  $\text{Var}(\ln(\mathbf{v})) = A \text{Var}(\mathbf{w}) A'$ .

In our preferred specifications, each of the two factors has two masspoints. We experimented with different numbers of masspoints for the latent factors. The maximum number that led to a significant improvement of the log likelihood function was three masspoints for each factor in the model with employment duration. However, the model fit in terms of a comparison between the original and simulated duration distributions differed only marginally from that based on two masspoints for each factor. Importantly, the model with two masspoints for each factor is computationally more attractive. Since we normalize the locations of the masspoints we can avoid evaluation of their derivatives. This speeds up computation by a factor of two to four, which is important when it comes to bootstrapping standard errors.

## 4.4 Estimation

Conditional on the observed covariates and the unobserved determinants, the joint density of the four durations for individual  $i$  is given by:

$$f_i = \prod_{j=1}^{N_i} \prod_{p=1}^{P_{ij}} \theta_p(t_{ijp} | \tau_{ijp}, x_{ij}(t_{ijp}), v_{ip})^{c_{ijp}} \exp\left[-\int_0^{t_{ijp}} \theta_p(r | \tau_{ijp}, x_{ij}(r), v_{ip}) dr\right],$$

where  $N_i$  is the number of cycles of person  $i$ ,  $P_{ij}$  is the total number of possible origin-destination state pairs associated with person  $i$ 's  $j$ th cycle, and  $c_{ijp}$  is a censoring indicator that equals one if the observation period of person  $i$  in cycle  $j$  ends with a transition from the origin state to the destination state indexed by  $p$ .  $\theta_p(\cdot)$  is the hazard rate that depends on elapsed duration in the current state  $t$ , lagged duration  $\tau$ , observed covariates  $x(t)$  and unobserved heterogeneity  $v$ . Since we allow for nonzero correlations of the unobserved heterogeneity terms in the different hazard rates, the likelihood function is not separable by individual and spell type. Thus, the individual likelihood contribution conditional on observed covariates and integrated

---

<sup>27</sup>All hazards contain an intercept.

over the vector of unobserved heterogeneity terms,  $\mathbf{v}$ , is:

$$L_i = \int \cdots \int \prod_{j=1}^{N_i} \prod_{p=1}^{P_{ij}} \theta_p(t_{ijp} | \tau_{ijp}, x_{ij}(t_{ijp}), v_{ip})^{c_{ijp}} \times \\ \exp \left[ - \int_0^{t_{ijp}} \theta_p(r_{ijp} | \tau_{ijp}, x_{ij}(r_{ijp}), v_{ip}) dr \right] dG(\mathbf{v}).$$

The final models involve the estimation of more than 350 and 500 parameters, respectively. The maximization routine uses a Newton-Raphson algorithm with analytic first and second derivatives.<sup>28</sup> To obtain standard errors, we implement the semi-parametric bootstrap procedure suggested in Gaure et al. (2010). For the coefficients on the observed covariates we draw values from their joint normal distribution.<sup>29</sup> Based on these draws we estimate a constrained model keeping the coefficients on the observed covariates fixed and maximizing over the coefficients associated with the unobserved heterogeneity distribution. We perform 250 bootstrap replications.<sup>30</sup> The full estimation results and the simulated model fit are displayed in sections F and G in the Online Appendix.

## 5 Causal Estimates of Training Impacts

### 5.1 Impacts of Training on the Treated

We use the parameter estimates to simulate the marginal distributions of treatment and nontreatment outcomes in the sample of treated persons. In particular, we consider all persons who receive training within the first 731 days of their first unemployment spell.<sup>31</sup> We exclude later training starts in order to avoid extrapolations beyond the time horizon of our data. We simulate ten unemployment-employment cycles for each person and treatment status.<sup>32</sup> From the simulated durations, we estimate the marginal distributions of treatment and nontreatment outcomes with the

---

<sup>28</sup>The maximum likelihood evaluator is implemented in Stata MP Version 11.1 and its matrix language Mata.

<sup>29</sup>Van der Vaart (1996) shows that, in the exponential hazard model with nonparametric unobserved heterogeneity, the coefficients attached to observed covariates are asymptotically normal.

<sup>30</sup>We use the bootstrap distribution of the unobserved heterogeneity parameters to test them for normality. The tests suggest no deviation from normality for 16 out of 17 parameters in model 1 (employment) and 7-9 parameters out of 13 in model 2 (earnings as second outcome).

<sup>31</sup>We obtain a subsample 7,046 of treated persons, 3,489 of which participate in short-term training and 3,557 in long-term training.

<sup>32</sup>Further details are provided in section H in the Online Appendix.

Kaplan and Meier (1958) estimator. We further assess training impacts through contrasting summary statistics of the marginal distributions, e.g. the truncated mean given in eq. (4). We compute standard errors for the summary statistics by re-running the simulation procedure 250 times using the parameter values from the bootstrap replications.<sup>33</sup>

Our data includes information on the planned and actually realized end dates of training programs. The planned enrolment length is assigned at the start of the program and may be shorter or longer than the realized duration.<sup>34</sup> We model duration dependence of treatment effects with dummy variables indicating different time intervals relative to the planned program end. This allows us to contrast in-program effects with post-program effects of training. Thus, we consider two measures for the impact of training on unemployment: the effect on the remaining time counted from the start of training and from the planned end of training. Panels (a) and (b) of figure 2 depict the simulated hazard rates with and without training from the program start onwards, panels (c) and (d) show the hazard rates starting from the planned program end.

— Insert figure 2 here. —

Panel (a) of figure 2 shows that participants in short-term training exit at a faster rate, compared to the situation without training, already from the beginning of the program onwards. Thus, a participation in short-term training, that has a median planned duration of four weeks, has no noticeable lock-in effect. To the contrary, it immediately helps participants in getting reemployed. The effect reaches its maximum about 65 days after program start. At that date, the hazard rate with short-term training is 18% higher than that without training. However, it then fades away relatively quickly. According to panel (c) of figure 2, the difference between the hazard rates with and without training is already very small three months after the planned program end.

The picture is rather different for long-term training programs that have a median planned duration of 201 days. In panel (b) of figure 2, we see that, during participation, people exit unemployment at a much lower rate compared to the situation of nonparticipation. Only after somewhat more than 200 days since program

---

<sup>33</sup>As the summary statistics on training impacts are smooth functionals of the survivor function we treat them as normal distributed when conducting hypothesis tests.

<sup>34</sup>Waller (2009) reports that about 20% of the persons who attend a training program for more than one week drop out early, i.e. before completing 80% of the scheduled enrolment length.



start, when the majority has completed training, the hazard rate recovers to a level slightly exceeding the one without participation. However, once we concentrate on the period following the scheduled program end, in panel (d) of figure 2, we see that long-term training has strong and persistent positive effects on the exit rate to employment. During the first three months after the planned program end, the hazard rate out of unemployment is about twice as high for the participants compared to the situation had they not participated. This effect slowly declines over time. One year after the planned program end, the hazard rate with participation in long-term training is about 40% higher than without training.

— Insert table 2 here. —

Table 2 provides summary statistics and bootstrapped standard errors on the impacts of short-term and long-term training on the remaining time in unemployment. In each table, the row labeled ‘NTO’ refers to the simulated nontreatment outcome and the row labeled ‘TT’ to the simulated treatment effect on the treated. The columns labeled ‘ $E[T|T \leq 731]$ ’ and ‘ $E[T|T \leq 365]$ ’ refer to the truncated means that are calculated according to eq. 4, now for the remaining time in unemployment until 731 days after the start of training and 365 days after the planned end, respectively. Counting from the program start, we see that short-term training reduces the expected remaining time in unemployment by a statistically significant 16 days, about 8.6%, whereas participating in long-term training significantly increases it by 95 days or 50.6%. Considering the time from the planned program end onwards, we see that long-term training has a stronger beneficial effect than short-term training: the latter reduces the average remaining duration in unemployment by 11 days (9.0%) whereas the reduction achieved with long-term training is 25 days (17.4%).

— Insert figure 3 here. —

Figure 3 displays the simulated impacts of training on the probability to stay employed beyond a given elapsed duration. For both training programs the probability to remain employed is higher with training than without at longer elapsed durations. The vertical difference between the survivors with and without training increases up to about 1.5 years after the beginning of the employment spell and remains constant thereafter. When the vertical difference stabilizes, the effect of long-term training is about twice as large as that of short-term training. Thus, long-term training increases the stability of employment much more strongly than short-term

training. Taking the horizontal difference between the survivor functions with and without training in each graph we obtain the quantile treatment effect. Figure 3 shows that the quantile treatment effects are increasing with decreasing percentiles of the survivor functions. The substantial heterogeneity of the quantile treatment effects suggests that there is also a considerable heterogeneity in the distribution of individual treatment effects.<sup>35</sup>

Table 3 translates the rather qualitative findings obtained from figure 3 into summary statistics and provides bootstrapped standard errors. The impact of short-term training on the expected employment duration (truncation point is 1825 days) is 7 days, 1.4%, and insignificant. The effect of long-term training on the mean employment duration is with 23 days or 5% somewhat bigger but still insignificant. However, at the median and the 70th percentile the horizontal distance between the cumulative distribution functions with and without long-term training is much larger and increasing. Similarly, the probability of survival in employment five years after the start of the spell is 9 percentage points or 45% higher with long-term training compared to the situation without. The corresponding number for short-term training is 5 percentage points, a bit less than 30%. 29% of the survival times in employment associated with long-term training exceed 1825 days compared to 22% of those with short-term training. This suggests, that the truncated mean underestimates in particular the effect of long-term training on the subsequent employment duration.

— Insert table 3 here. —

Table 4 shows the impacts of training on average daily earnings during the first year after reemployment. Short-term training has no effect on earnings. The point estimate is almost zero and not significant statistically. Contrary to the descriptive findings, long-term training leads to significant earnings gains of €3.37 or 7% a day. These estimates likely underestimate the full monetary return to training as our earnings measure refers to the first year after reemployment only. In fact, our findings for employment stability suggest that especially persons treated with long-term training accumulate more work experience than without training over the longer run. This may lead to additional earnings gains in the medium and long run.<sup>36</sup>

---

<sup>35</sup>See Heckman, Smith, and Clements (1997) for a discussion of properties of the distribution of individual treatment effects that can be inferred from the marginal distributions of treatment and nontreatment outcomes.

<sup>36</sup>To get an idea of the persistence of training impacts in later unemployment and employment

— Insert table 4 here. —

## 5.2 Importance of the Timing of Training

We model the time dependence of treatment effects as proportional shifts of the hazard rate out of unemployment that vary with time relative to the planned program end. In addition, we include a linear and a quadratic term of the log unemployment duration at program start in the hazard rate. This specification allows for complex effects of the timing of training during the unemployment spell. In order to assess the *ceteris paribus* effect of changing the starting date of training during unemployment, we predict the truncated mean unemployment duration associated with different starting dates for a reference person with characteristics at the mean or mode of each covariate. Specifically, this person is a male German aged 38 at the start of the spell, married without children, living in the German federal state of North Rhine-Westphalia, holding a secondary schooling degree reached at the end of the 10th grade (Realschule) and a vocational training degree, previously employed as a bluecollar craftsman in the trade and transport sector, with log daily earnings of 3.9 (i.e. in the third quartile), entitled to 340 days of unemployment benefits, considered as having relevant vocational qualification and experience by the caseworker, and starting his unemployment spell in the first quarter of 2000.<sup>37</sup> The planned program durations are set to their medians at 26 (short-term training) and 201 days (long-term training), respectively. The unobserved heterogeneity terms are set to their mean values.

— Insert figure 4 here. —

Panel (a) of figure 4 shows the impact of the timing of training on the truncated expected unemployment duration measured from the start of unemployment until day 1825 and the right panel the impact on the truncated expected remaining unemployment duration measured from the start of training until day 731 after training start.<sup>38</sup> Panel (a) suggests that, in absolute value, the impact of both short-term and

---

spells, we include indicators for lagged training participation. Overall, the estimates indicate some persistence across spells. However, positive effects on employment stability tend to be offset by negative effects on unemployment duration.

<sup>37</sup>The original estimation includes time-varying dummies indicating different seasons of the year. In the simulation, we assign each of these variables a time-constant value representing their share in a given calendar year. This way we obtain expected outcome durations that are seasonally adjusted.

<sup>38</sup>The truncated expectations are calculated according to eq. (4).

long-term training on the total unemployment duration is strongest when training is started early during unemployment, but with opposite signs. A participation in long-term training starting during the first three months of unemployment increases the total expected time in unemployment by up to 151 days, whereas a participation in short-term training in the same period decreases the total expected time in unemployment by up to 43 days. With increasing elapsed unemployment duration at program start the impacts of both training programs on the total expected unemployment duration decrease in absolute value. The effect of long-term training converges to +16 days whereas that of short-term training approaches to zero as the time in unemployment at program start increases.

Panel (b) of figure 4 displays the impact of the timing of training on the remaining time in unemployment counted from the start of the program. The impact of short-term training is largely constant at about  $-40$  days regardless of the starting date during unemployment. In contrast, the impact of long-term training on the remaining time in unemployment displays a declining pattern, from about  $+115$  days for program starts during the first three months to  $+60$  days for program starts in the eighth quarter after the beginning of unemployment. The different patterns in panel (a) and (b) of figure 4 reflect that longer survival times are weighted differently in the two graphs. While the expected value from the start of training (panel (b)) conditions on survival in unemployment until program start, the expected value from the start of unemployment (panel (a)) accounts for the possibility that the unemployment spell may terminate before the start of training. Therefore, in panel (a), training impacts at later points in time during unemployment receive a lower weight than training impacts occurring at the start of unemployment. This difference in weighting largely drives the apparently diminishing effectiveness of short-term training in panel (a), while the analogue does not hold for long-term training. In fact, our findings suggest that long-term training programs starting later during unemployment lead to a considerably smaller increase in unemployment duration.

### 5.3 Effect Heterogeneity Across Observed Characteristics

Our estimated models flexibly account for heterogeneous treatment effects across observed and unobserved variables. To examine the differences across observed characteristics, we aggregate the simulated treatment and nontreatment outcomes in a given subgroup, e.g. men, and calculate the treatment effect in this subgroup. The entries in tables 5 and 6 show the program impact on the treated for the subgroup given in the row label. The treatment effects refer to the truncated means

for the outcomes remaining time in unemployment from program start and planned program end as well as daily earnings. The truncation points are 731 days, 365 days and €144, respectively. For the outcome employment duration, we present results on the quantile treatment effect evaluated at the median and on the probability to remain employed after 1825 days. The last row for each set of subgroups provides the  $p$ -value of a Wald test that the treatment effects are equal across subgroups.<sup>39</sup>

— Insert table 5 here. —

— Insert table 6 here. —

Several interesting differences across groups emerge. Female participants in long-term training tend to benefit more than males, whereas no consistent gender differences exist for short-term training. Long-term training increases the remaining time in unemployment measured from the program start by 87 days (44.6%) for women and 102 days (56.8%) for men. Female participants in long-term training achieve higher employment gains of 291 days (64.8%) at the median and earnings gains of €4.37 (10.4%) on average per calendar day as opposed to 115 days (36.7%) and €2.69 (5.2%), respectively, for males. For older workers aged 45 and above, short-term training tends to work less well than for the younger age groups. However, the null hypothesis of equal of treatment effects across age groups cannot be rejected for any of the impact measures at conventional significance levels. Comparing the results for foreign nationals to those of Germans participating in long-term training, we find very small effects on employment duration for foreign people but large effects for Germans (e.g. a quantile treatment effect of +312 days, 84.7%). However, foreign nationals participating in long-term training exhibit higher earnings gains of €4.83 (11.1%) on average per calendar day as opposed to €2.80 (5.8%).

High-skilled participants in long-term training holding a university degree or previously working as senior officials, managers and professionals exhibit very strong lock-in effects, i.e. a 107 to 140 days (around 60%) increase in the remaining unemployment duration measured from the program start. For participants in long-term training previously working in low- to medium-skilled occupations the corresponding figure lies between +90 and +100 days (around 50%). Persons with low formal qualifications as well as persons previously working in elementary occupations exhibit

---

<sup>39</sup>As an alternative, in section I in the Online Appendix, we provide evidence from hypothesis tests that indicate whether the proportionate treatment effect on the hazard rate is different for a particular subgroup keeping everything else constant.

small employment gains: at the median the effect is +49 days (16.2%) for low skilled and +60 days (20.1%) for elementary occupations, respectively. In contrast, people with medium formal qualifications and people previously working in medium-skilled job that are intensive in analytic and interactive tasks show very strong employment gains. Technicians and associate professionals, for instance, gain 432 employment days (89.8%) at the median and have a 16.6 percentage points (69.3%) higher probability to remain employed after five years. High skilled and persons in high-skilled occupations also show big point estimates for employment duration, but the standard errors tend to be large as well.

When it comes to earnings impacts of long-term training the ordering tends to be reversed. Persons with low formal education and persons previously working in low- and medium-skilled manual jobs (in particular, elementary occupations and service workers) reap the highest earnings gains of €4.12, €5.53 and €6.80, respectively. In relative terms, these earnings gains are in the order of 10.3%, 15.1%, and 18.9%, respectively, a substantial amount from an economic point of view. Senior officials and managers, experience an earnings gain of €8.60 per calendar day, which corresponds to a gain of 11.9% in relative terms. The earnings gains for persons holding a university degree as well as for people previously working in high- and medium-skilled occupations that are intensive in analytic and interactive tasks, in contrast, are small (between €0.33 and 1.85).

The differing impacts of long-term training across occupation groups may also contribute to explaining the differences across gender. In our sample, 36.8% of female job-seekers previously worked as clerk or technician and associate professional as opposed to only 18.1% of the males. These occupations reap the highest employment gains. 23.9% of the women as opposed to 7.5% of the men previously worked as service workers, a group with above average earnings gains. 49% of the male unemployed previously worked in production occupations (plant/machine operator or craft and related occupations) that show below average employment gains from long-term training. However, the earnings gains for these occupations tend to be above average. Similarly, the smaller employment and higher earnings effects for foreign nationals as opposed to Germans participating in long-term training may be associated with their higher shares in elementary and production occupations.

As regards short-term training, the group of technicians and associate professionals exhibits exceptionally high benefits in terms of employment stability with a quantile treatment effect of +129 days (25.5%). For the other occupation groups the quantile treatment effects for employment duration are smaller. However, we find sizeable

impacts of 5 to 8 percentage points on the probability to be still employed after five years for five out of nine occupation groups. Overall, the differences across skill and occupation groups do not suggest a coherent interpretation.

In sum, these patterns suggest that long-term training programs facilitate the occupational advancement of heterogeneous groups of participants. However, their beneficial effects manifest in different ways across different groups. While they have large positive effects on employment stability for people in medium- to high-skilled analytic and interactive jobs, there are almost no employment effects for people in low- to medium-skilled service and production occupations. However, the latter groups show above average earnings gains, whereas the former earn hardly more than without training. The often sizeable employment and earnings gains come at the cost of prolonged unemployment spells. The opportunity cost of participating in long-term training is highest for people holding a university degree as well as for senior officials and managers. Short-term training programs do not harm any of the groups considered but their beneficial effects are generally limited in magnitude.

## 5.4 Cost-Benefit Assessment

From a policy point of view it is important to know whether the gains of participating in short-term or long-term training outweigh the costs. In order to get an idea of the cost-effectiveness of the programs consider the following back-of-the-envelope calculations. A short-term training course costs on average €560 and a long-term training course €5,850 (cf. table B1 in the Online Appendix). The employment agency pays on average €1,050 per month for an unemployed entitled to unemployment benefits. Extrapolating beyond the limits imposed by our data we can calculate the impacts of training on the nontruncated means of unemployment and employment duration as well as earnings as averages across our sample of treated persons.<sup>40</sup> A participation in short-term training reduces the expected unemployment duration by 53 days (-7.5%), increases the expected employment duration by 162 days (+19.1%) and reduces average daily earnings by €-0.46 (-1.0%). Long-term training increases unemployment duration by 97 days (+15.7%), employment duration by 330 days (+34.9%) and daily earnings by €3.50 (+7.2%).

---

<sup>40</sup>We compute the (nontruncated) expected treatment and nontreatment outcomes as follows:  $E[T] = \int_0^{\bar{t}} S(r)dr + S(\bar{t})/\theta(\bar{t})$ , with  $S(t)$  the survivor function and  $\theta(t)$  the hazard rate at  $T = t$ . This means, we extrapolate beyond truncation point  $\bar{t}$  assuming a constant hazard rate equal to the hazard rate at  $\bar{t}$ . We set  $\bar{t}$  to 1460 days for unemployment duration, 1825 days for employment duration, and €144 for daily earnings.

These calculations indicate that short-term training is likely cost effective, both from the taxpayer's and the participant's perspective. It reduces unemployment duration and the associated transfer payments by about €1,850. These savings in transfer payments exceed the course fees by a factor of three. In addition, short-term training positively affects the expected time in subsequent employment. In contrast, the picture is less clear for long-term training. With the additional benefit payments arising during the extra time in unemployment, a long-term training course causes about €9,000 of additional costs for the taxpayer compared to the situation of nonparticipation. Yet in view of the substantial positive effects on employment duration and the moderate earnings gains, it might be possible that in the long run also long-term training pays off from a fiscal point of view. As the benefits in terms of reduced unemployment insurance payments and higher tax revenues accumulate over time they may eventually exceed the initial investment of €9,000. Further, from the perspective of the participant, long-term training programs well seem to be attractive on average. Indeed, while unemployed participants receive relatively generous unemployment benefits. Once they have completed their long-term training course, on average, they quickly take up a job that is more stable and pays more than without training. A full assessment of the cost-effectiveness of both training programs according to a social cost-benefit criterion, that considers the costs and benefits that accrue to taxpayers and participants alike, would require further assumptions and information, e.g. on discount rates and tax scales.

## 6 Conclusion

This study investigates and compares the dynamic causal effects of short-term, job-search oriented training and long-term, human capital intensive training schemes. Our empirical analysis uses rich administrative data for Germany, where both program types are used at the same time. We examine and compare the separate effects of both programs on unemployment and employment spells as well as daily earnings, taking into account the heterogeneity of training impacts according to the timing of participation during unemployment as well as across different subgroups of participants.

We find that participating in short-term training reduces the remaining time in unemployment and has moderate positive effects on subsequent job stability. Long-term training programs initially prolong the remaining time in unemployment, but once the scheduled program end is reached participants exit to employment at a



much faster rate than without training. Moreover, participants in long-term training benefit from substantially more stable employment spells and higher earnings.

Importantly, our findings point to the possibility of improving the efficiency of long-term training programs through a careful targeting. Specifically, we find that the opportunity cost of participating in terms of prolonged unemployment is lower for people with lower chances to exit unemployment on their own, e.g. long-term unemployed and low skilled. Persons without formal education degree and persons previously working in low- and medium-skilled manual occupations reap particularly high benefits with respect to earnings. Persons previously working in medium-skilled analytic and interactive occupations achieve substantial gains in terms of employment stability. Thus, long-term training programs seem well effective in supporting the occupational advancement of very heterogeneous groups of people, including those with generally weak labor market prospects. However, from a more narrow, fiscal perspective only the low-cost short-term training schemes are on average cost efficient in the short run.<sup>41</sup> In light of the recent policy shifts favoring short, job-search oriented training programs a more balanced use of both types of training seems warranted. In fact, while the shorter programs may appeal for their cost efficiency the longer programs seem more effective in tackling structural deficits.

From a conceptual point of view, our study highlights that time is an important dimension of program evaluation not appropriately accounted for in conventional static and experimental research designs. A detailed analysis of the dynamics of program assignment and labor market outcomes, considering also impact heterogeneity allows a deeper understanding of how programs work. This knowledge is important for an optimal use of public employment services.

## References

- Abbring, Jaap H. (2007). "The Event-History Approach to Program Evaluation." Tinbergen Institute Discussion Paper 2006-57/3.
- Abbring, Jaap H. and James J. Heckman (2007). "Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation." In James J. Heckman and Edward E. Leamer (eds), *Handbook of Econometrics*,

---

<sup>41</sup>All these conclusions hold with the caveat that general equilibrium effects are not accounted for.

Vol. 6B. North-Holland, Amsterdam, 5145-5303.

- Abbring, Jaap H. and Gerard J. van den Berg (2003a). "The Nonparametric Identification of Treatment Effects in Duration Models." *Econometrica*, 71 (5), 1491-1517.
- Abbring, Jaap H. and Gerard J. van den Berg (2003b). "The Identifiability of the Mixed Proportional Hazards Competing Risks Model." *Journal of the Royal Statistical Society, Series B*, 65 (3), 701-710.
- Abbring, Jaap H. and Gerard J. van den Berg (2004). "Analyzing the Effect of Dynamically Assigned Treatments Using Duration Models, Binary Treatment Models, and Panel Data Models." *Empirical Economics*, 29 (1), 5-20.
- Abbring, Jaap H. and Gerard J. van den Berg (2005). "Social Experiments and Instrumental Variables with Duration Outcomes." Tinbergen Institute Discussion Paper TI 2005-047/3.
- Acemoglu, Daron and David Autor (2011). "Skills, Tasks and Technologies: Implications for Employment and Earnings." In Orley Ashenfelter and David Card (eds), *Handbook of Labor Economics*, Vol. 4B, 1043-1166.
- Ashenfelter, Orley (1978). "Estimating the Effect of Training Programs on Earnings." *Review of Economics and Statistics*, 60 (1), 47-57.
- Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2008). "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics*, 90 (2), 300-323.
- Bergemann, Annette, Bernd Fitzenberger, and Stefan Speckesser (2009). "Evaluating the Dynamic Employment Effects of Training Programs in East Germany Using Conditional Difference-in-Differences." *Journal of Applied Econometrics*, 24 (5), 797-823.
- Blank, Rebecca (2002). "Evaluating Welfare Reform in the United States." *Journal of Economic Literature*, 40 (4), 1105-1166.
- Bonnal, Liliane, Denis Fougère, and Anne Sérandon (1997). "Evaluating the Impact of French Employment Policies on Individual Labour Market Histories." *Review of Economic Studies*, 64 (4), 683-713.
- Brinch, Christian (2007). "Nonparametric Identification of the Mixed Hazards Model with Time-Varying Covariates." *Econometric Theory*, 23 (2), 349-354.

- Card, David, Jochen Kluve, and Andrea Weber (2010). "Active Labour Market Policy Evaluations: A Meta-Analysis." *Economic Journal*, 120 (548), F452-F477.
- Couch, Kenneth A. (1992). "New Evidence on the Long-Term Effects of Employment Training Programs." *Journal of Labor Economics*, 10 (4), 380-388.
- Crépon, Bruno, Muriel Dejemeppe, and Marc Gurgand (2005). "Counseling the Unemployed: Does It Lower Unemployment Duration and Recurrence?" IZA Discussion Paper No. 1796.
- Crépon, Bruno, Marc Ferracci, and Denis Fougère (2011). "Training the Unemployed in France: How Does It Affect Unemployment Duration and Recurrence?" forthcoming *Annales d'Economie et de Statistique*.
- Crépon, Bruno, Marc Ferracci, Grégory Jolivet and Gerard J. van den Berg (2009). "Active Labor Market Policy Effects in a Dynamic Setting." *Journal of the European Economic Association*, 7 (2-3), 595-605.
- de Luna, Xavier and Per Johansson (2009). "Non-Parametric Inference for the Effect of a Treatment on Survival Times with Application in the Health and Social Sciences." IZA Discussion Paper No. 3966.
- Donald, Stephen G., David A. Green, and Harry J. Paarsch (2000). "Differences in Wage Distributions between Canada and the United States: An Application of a Flexible Estimator of Distribution Functions in the Presence of Covariates." *Review of Economic Studies*, 67 (4), 609-633.
- Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg (2009). "Revisiting the German Wage Structure." *Quarterly Journal of Economics*, 124 (2), 843-881.
- Dyke, Andrew, Carolyn J. Heinrich, Peter R. Mueser, Kenneth R. Troske, and Kyung-Seong Jeon (2006). "The Effects of Welfare-to-Work Program Activities on Labor Market Outcomes." *Journal of Labor Economics*, 24 (3), 567-607.
- Eberwein, Curtis, John C. Ham, and Robert J. LaLonde (1997). "The Impact of Being Offered and Receiving Classroom Training on the Employment Histories of Disadvantaged Women: Evidence from Experimental Data." *Review of Economic Studies*, 64 (4), 655-682.
- Elsby, Michael W.L., Bart Hobijn, and Ayşegül Şahin (2010). "The Labor Market in the Great Recession." *Brookings Papers on Economic Activity*, Spring, 1-69.

- European Commission (2002). "Taking Stock of Five Years of the European Employment Strategy." COM 416(01), Commission of the European Communities, Brussels.
- Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter (2008). "Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany." *Annales d'Economie et de Statistique*, 91-92, 321-355.
- Fredriksson, Peter and Per Johansson (2008). "Dynamic Treatment Assignment: The Consequences for Evaluations Using Observational Data." *Journal of Business and Economic Statistics*, 26 (4), 435-445.
- Gaure, Simen, Knut Røed, Gerard J. van den Berg and Tao Zhang (2010). "Estimation of Heterogeneous Treatment Effects on Hazard Rates." IZA Discussion Paper No. 4794.
- Gaure, Simen, Knut Røed, and Tao Zhang (2007). "Time and Causality: A Monte Carlo Assessment of the Timing-of-Events Approach." *Journal of Econometrics*, 141 (2), 1159-1195.
- Greenberg, David H., Charles Michalopoulos, and Philip K. Robins (2003). "A Meta-Analysis of Government Sponsored Training Programs." *Industrial and Labor Relations Review*, 57 (1), 31-53.
- Gritz, Marc (1993). "The Impact of Training on the Frequency and Duration of Employment." *Journal of Econometrics*, 57 (1-3), 21-51.
- Ham, John C., and Robert J. LaLonde (1996). "The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training." *Econometrica*, 64 (1), 175-205.
- Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith (1999). "The Economics and Econometrics of Active Labor Market Programs." In Orley Ashenfelter and David Card (eds), *Handbook of Labor Economics*, Vol. 3A, Amsterdam: Elsevier Science, 1865-2097.
- Heckman, James J. and Burton Singer (1984). "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica*, 52 (2), 271-320.

- Heckman, James J. and Jeffrey A. Smith (1999). "The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies." *Economic Journal*, 109 (457), 313-348.
- Heckman, James J., Hidehiko Ichimura, and Petra E. Todd (1997). "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies*, 64 (4), 605-654.
- Heckman, James J., Jeffrey Smith, and Nancy Clements (1997). "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts." *Review of Economic Studies*, 64 (4), 487-535.
- Honoré, Bo E. (1993). "Identification Results for Duration Models with Multiple Spells." *Review of Economic Studies*, 60 (1), 241-246.
- Horný, Guillaume and Matteo Picchio (2010). "Identification of Lagged Duration Dependence in Multi-Spell Competing Risks Models." *Economics Letters*, 106 (3), 241-243.
- Hotz, V. Joseph, Guido W. Imbens and Jacob A. Klerman (2006). "Evaluating the Differential Effects of Alternative Welfare-to-Work Training Components: A Reanalysis of the California GAIN Program." *Journal of Labor Economics*, 24 (3), 521-566.
- Hujer, Reinhard, Stephan L. Thomsen and Christopher Zeiss (2006a): "The Effects of Short-Term Training Measures on the Individual Unemployment Duration in West Germany." ZEW Discussion Paper 06-65.
- Hujer, Reinhard, Stephan L. Thomsen and Christopher Zeiss (2006b): "The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany." *AStA Advances in Statistical Analysis*, 90 (2), 299-321.
- Kaplan, Edward L. and Paul Meier (1958). "Nonparametric Estimation from Incomplete Observations." *Journal of the American Statistical Association*, 53 (282), 457-481.
- Kluve, Jochen (2010). "The Effectiveness of European Active Labor Market Policy." *Labour Economics*, 17 (6), 904-918.

- Lalive, Rafael, Jan C. van Ours and Josef Zweimüller (2008). "The Impact of Active Labour Market Programmes on the Duration of Unemployment in Switzerland." *Economic Journal*, 118 (1), 235-257.
- Lechner, Michael, Ruth Miquel, and Conny Wunsch (2011). "Long-Run Effects of Public Sector Sponsored Training in West Germany." *Journal of the European Economic Association*, 9 (4), 742-784.
- Lok, Judith (2008). "Statistical Modeling of Causal Effects in Continuous Time." *Annals of Statistics*, 36 (3), 1464-1507.
- Neyman, Jerzy (1923), edited and translated by Dorota M. Dabrowska and Terrence P. Speed (1990). "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9." *Statistical Science*, 5 (4), 465-472.
- OECD (2005, 2006). *OECD Employment Outlook*. Paris (various issues).
- Richardson, Katarina and Gerard J. van den Berg (2006). "Swedish Labor Market Training and the Duration of Unemployment." IZA Discussion Paper No. 2314.
- Robins, James M. (1998). "Structural Nested Failure Time Models." In Peter Armitage and Theodore Colton (eds), *The Encyclopedia of Biostatistics*, Chichester: John Wiley and Sons, 4372-4389.
- Roy, Andrew D. (1951). "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers*, 3 (2), 135-146.
- Rubin, Donald B. (1974). "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology*, 66 (5), 688-701.
- Sianesi, Barbara (2004). "An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s." *Review of Economics and Statistics*, 86 (1), 133-155.
- Sianesi, Barbara (2008). "Differential Effects of Active Labour Market Programs for the Unemployed." *Labour Economics*, 15 (3), 370-399.
- Spitz-Oener, Alexandra (2006). "Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure." *Journal of Labor Economics*, 24 (2), 235-270.

Waller, Marie (2009). “Many Dropouts? Never Mind! – Employment Prospects of Dropouts from Training Programs.” CDSE Mannheim Discussion Paper No. 63.

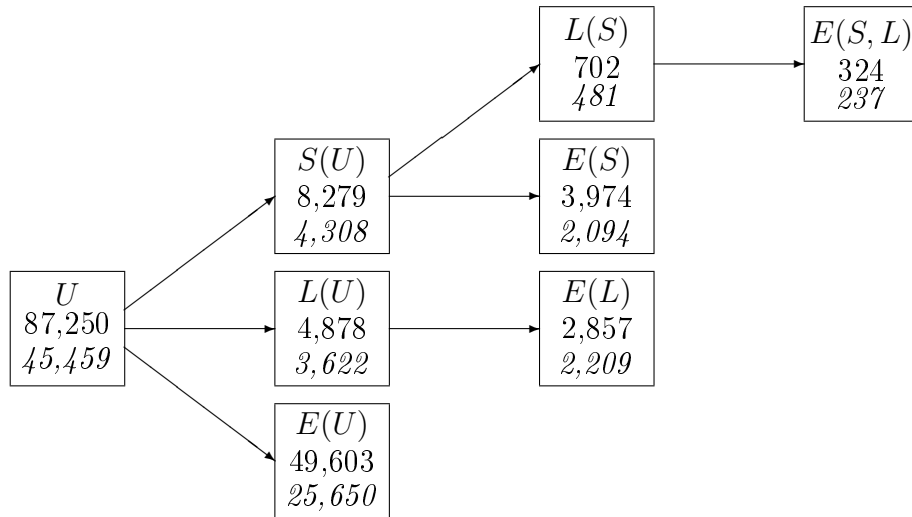
Weber, Andrea, and Helmut Hofer (2004a). “Are Job Search Programs a Promising Tool? A Microeconometric Evaluation for Austria.” IZA Discussion Paper No. 1075.

Weber, Andrea, and Helmut Hofer (2004b). “Employment Effects of Early Interventions on Job Search Programs.” IZA Discussion Paper No. 1076.

van der Vaart, Aad (1996). “Efficient Maximum Likelihood Estimation in Semiparametric Mixture Models.” *Annals of Statistics*, 24 (2), 862-878.

## Figures

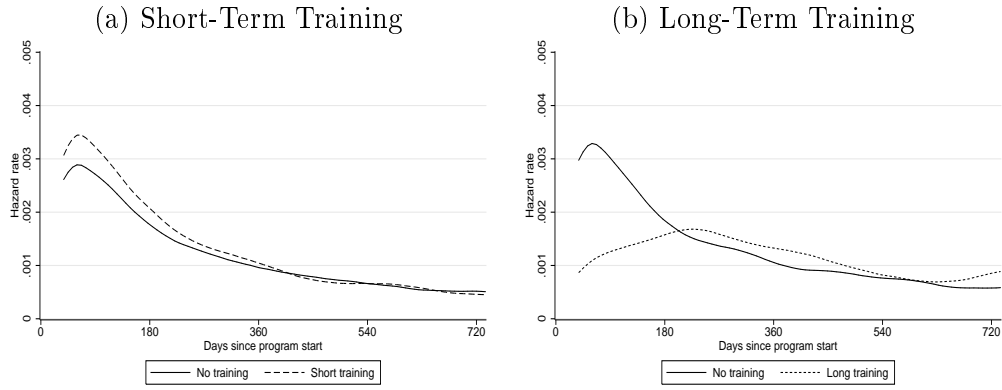
Figure 1: Number of Observations at Different Phases of the Unemployment-Employment Cycle



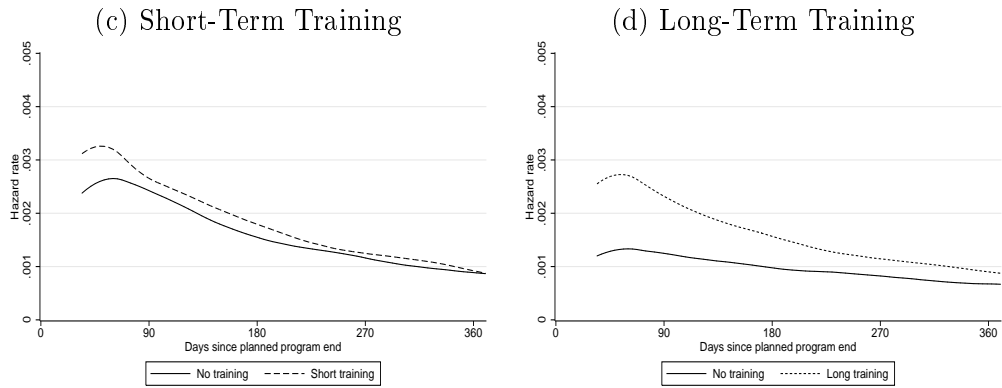
Notes:  $U, S, L, E$  denote the labor market states open unemployment, unemployment treated with short-term training, unemployment treated with long-term training and employment, respectively. The states in parentheses indicate dependence of current states on the past trajectory. The first number in each box refers to the number of observations using repeated cycles per person, the second number (in italics) to the number of observations using only the first cycle.

Figure 2: Impacts of Training on the Hazard Rate out of Unemployment

Hazard Rate out of Unemployment from Program Start



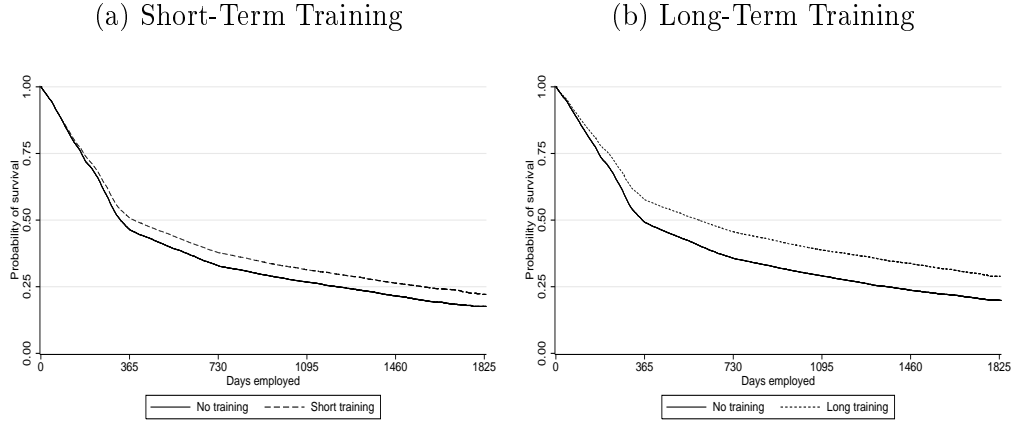
Hazard Rate out of Unemployment from Planned Program End



Notes: Kaplan-Meier (1958) estimates of the hazard rates are based on simulated unemployment durations with and without training for a subset of the originally treated persons, cf. p. 22. The bandwidth for the kernel smooth of the hazard rates is 30 days. Estimates within 30 days of the left boundary are omitted.

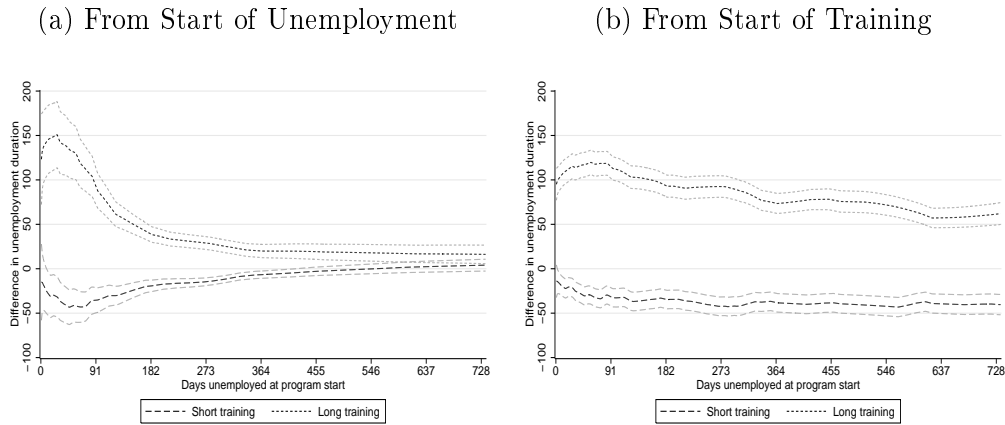


Figure 3: Impacts of Training on the Probability of Survival in Employment



Notes: Kaplan-Meier (1958) estimates of the probability to remain employed are based on simulated employment durations with and without training for a subset of the originally treated persons, cf. p. 22.

Figure 4: Training Impacts on Unemployment Duration by Starting Date



Notes: Calculations are based on the predicted means measured from the start of the spell (panel (a)) and from the start of training (panel (b)), respectively, for a reference person, cf. p. 26. The truncated means are calculated according to eq. (4). The truncation point is 1825 days in the left panel and 731 days in the right. The grey lines are 95% confidence intervals based on bootstrapped standard errors with 250 replications.

## Tables

Table 1: Unemployment and Employment Outcomes by Treatment Status

	Unemployment		Employment		Earnings	
	$E(T T \leq 1825)$	Median	$E(T T \leq 1825)$	Median	$E(T T \leq 144)$	Median
No training	219	160	375	284	49.87	49.25
Short training	235	172	361	304	45.08	43.29
Long training	570	607	477	506	50.16	50.13
All	261	167	380	287	49.56	48.87
Observations	87,250		56,758		30,190	

Notes: Unemployment and employment durations are measured in days, daily earnings in € (real values, reference year 2000).

Table 2: Nontreatment Outcomes (NTO) and Treatment Effects on the Treated (TT) for Remaining Unemployment Duration

(a) Remaining Time in Unemployment from Program Start					
	$E(T T \leq 731)$	$\Pr(T > 731)$	P15	P30	P45
Short-term training					
NTO	191.39 (2.20)	0.39 (0.01)	54 (1.44)	122 (4.01)	264 (11.39)
TT	-16.37 (4.02)***	-0.04 (0.02)*	-7 (2.83)**	-21 (6.99)***	-59 (21.02)***
Long-term training					
NTO	187.68 (1.81)	0.35 (0.01)	46 (1.24)	106 (2.92)	225 (8.52)
TT	95.01 (4.19)***	0.08 (0.03)***	97 (8.61)***	154 (15.19)***	211 (37.66)***
(b) Remaining Time in Unemployment from Planned Program End					
	$E(T T \leq 365)$	$\Pr(T > 365)$	P15	P30	P45
Short-term training					
NTO	121.54 (0.84)	0.53 (0.01)	55 (1.52)	135 (4.65)	320 (14.00)
TT	-10.84 (1.57)***	-0.06 (0.02)***	-23 (3.12)***	-35 (8.81)***	-94 (24.89)***
Long-term training					
NTO	143.64 (1.41)	0.69 (0.01)	118 (4.29)	332 (13.37)	—
TT	-24.54 (2.97)***	-0.17 (0.03)***	-69 (6.13)***	-201 (19.35)***	—

Notes: Unemployment durations are measured in days. P15, P30, P45 refer to the percentiles of the cumulative distribution function. Calculations are based on simulated unemployment durations with and without training for a subset of the originally treated persons. Bootstrapped standard errors based on 250 replications in parentheses. \*, \*\* and \*\*\* denote significance at the 10%-, 5%- and 1%- level, respectively.

Table 3: Nontreatment Outcomes (NTO) and Treatment Effects on the Treated (TT) for Employment Duration

	$E(T T \leq 1825)$	$\Pr(T > 1825)$	P30	P50	P70
Short-term training					
NTO	435.57 (9.07)	0.17 (0.01)	207 (4.72)	330 (7.20)	869 (38.67)
TT	6.99 (14.31)	0.05 (0.02)***	11 (11.22)	48 (40.20)	279 (129.57)**
Long-term training					
NTO	445.96 (7.17)	0.20 (0.01)	216 (4.03)	354 (8.50)	1050 (47.70)
TT	22.71 (16.61)	0.09 (0.03)***	35 (15.20)**	215 (92.31)**	682 (176.65)***

Notes: Employment duration is measured in days. P30, P50, P70 refer to the percentiles of the cumulative distribution function. Calculations are based on simulated employment durations with and without training for a subset of the originally treated persons. Bootstrapped standard errors based on 250 replications in parentheses. \*, \*\* and \*\*\* denote significance at the 10%-, 5%- and 1%- level, respectively.

Table 4: Nontreatment Outcomes (NTO) and Treatment Effects on the Treated (TT) for Daily Earnings

	$E(T T \leq 144)$	$\Pr(T > 144)$	P25	P50	P75
Short-term training					
NTO	46.29 (0.83)	0.015 (0.001)	26.22 (0.66)	45.54 (1.06)	63.55 (1.03)
TT	-0.05 (0.58)	-0.003 (0.001)***	0.20 (0.45)	-0.06 (0.72)	-0.62 (0.70)
Long-term training					
NTO	47.02 (0.88)	0.014 (0.001)	26.58 (0.70)	46.67 (1.12)	64.67 (1.12)
TT	3.37 (1.48)**	0.001 (0.003)	3.59 (1.33)***	4.58 (1.69)***	3.56 (2.05)*

Notes: Earnings per calendar day are measured in € (real values, reference year 2000). P25, P50, P75 refer to the percentiles of the cumulative distribution function. Calculations are based on simulated earnings with and without training for a subset of the originally treated persons. Bootstrapped standard errors based on 250 replications in parentheses. \*, \*\* and \*\*\* denote significance at the 10%-, 5%- and 1%- level, respectively.

Table 5: Heterogeneous Treatment Effects of Short-Term Training

		Remaining Days in Unemployment from				
		Program Start	Planned Prog. End	Employment		Earnings
		$E[T T \leq 731]$	$E[T T \leq 365]$	P50	$\Pr(T > 1825)$	$E[T T \leq 144]$
Gender	Male	-20.19 (4.56)	-11.43 (1.88)	30 (23.80)	0.053 (0.021)	-0.91 (0.68)
	Female	-11.58 (5.28)	-10.13 (1.71)	92 (56.97)	0.043 (0.020)	1.24 (0.78)
	<i>p</i> -Value of Wald Test of Equality	0.123	0.474	0.140	0.588	0.014
Nationality	German	-15.22 (4.20)	-10.39 (1.70)	61 (51.82)	0.042 (0.019)	-0.72 (0.61)
	Foreign	-18.94 (4.36)	-11.81 (1.84)	52 (30.38)	0.062 (0.022)	1.49 (0.92)
	<i>p</i> -Value of Wald Test of Equality	0.190	0.387	0.800	0.276	0.018
Age	Age 25-34	-17.72 (4.38)	-12.13 (1.89)	36 (30.83)	0.039 (0.019)	0.04 (0.60)
	Age 35-44	-17.55 (4.47)	-10.51 (2.00)	64 (47.29)	0.052 (0.018)	-0.12 (0.58)
	Age 45+	-10.87 (4.65)	-8.57 (1.94)	64 (45.28)	0.064 (0.027)	-0.08 (0.65)
	<i>p</i> -Value of Wald Test of Equality	0.104	0.202	0.301	0.424	0.600
Education	No voc. degree and missing	-15.23 (4.12)	-11.57 (1.69)	17 (14.15)	0.014 (0.017)	0.52 (0.61)
	Vocational training degree	-18.05 (4.37)	-11.24 (1.73)	102 (54.29)	0.072 (0.021)	-0.17 (0.61)
	University degree	-7.81 (5.48)	-3.49 (3.35)	104 (57.74)	0.011 (0.028)	-0.16 (0.74)
	<i>p</i> -Value of Wald Test of Equality	0.099	0.025	0.115	0.000	0.088
Occupation	Elementary occupations	-12.72 (4.10)	-9.45 (2.37)	19 (17.18)	0.065 (0.022)	0.46 (0.65)
	Agricultural occ. and missing	15.12 (10.85)	0.67 (5.56)	10 (17.74)	-0.026 (0.052)	0.12 (0.83)
	Plant/machine operator	-16.51 (4.24)	-13.42 (2.84)	31 (26.07)	0.018 (0.030)	0.10 (0.72)
	Craft and related occupations	-23.85 (5.32)	-13.84 (2.11)	16 (11.35)	0.045 (0.022)	-0.46 (0.73)
	Service worker	-14.47 (5.26)	-9.03 (2.71)	73 (49.19)	0.054 (0.027)	1.53 (0.72)
	Clerk	-17.11 (5.55)	-10.21 (2.09)	99 (64.55)	0.065 (0.026)	-0.38 (0.96)
	Techn., assoc. professional	-15.92 (5.79)	-16.57 (2.41)	129 (63.75)	0.077 (0.032)	-0.19 (0.86)
	Professional	-10.99 (6.69)	-1.55 (3.17)	44 (63.59)	0.029 (0.025)	-0.38 (0.79)
	Senior official, manager	-29.48 (10.02)	-8.85 (4.62)	105 (78.81)	0.028 (0.051)	0.62 (1.09)
	<i>p</i> -Value of Wald Test of Equality	0.014	0.000	0.008	0.324	0.034

Notes: Unemployment and employment durations are measured in days, earnings per calendar day in € (real values, reference year 2000). P50 refers to the difference between treatment and nontreatment outcomes at the median of the cumulative distribution functions. Calculations are based on simulated outcomes with and without training for a subset of the originally treated persons. Bootstrapped standard errors based on 250 replications in parentheses.

Table 6: Heterogeneous Treatment Effects of Long-Term Training

		Remaining Days in Unemployment from				
		Program Start	Planned Prog. End	Employment		Earnings
		$E[T T \leq 731]$	$E[T T \leq 365]$	P50	$\Pr(T > 1825)$	$E[T T \leq 144]$
Gender	Male	102.22 (4.27)	-24.19 (3.32)	115 (74.82)	0.077 (0.030)	2.69 (1.57)
	Female	87.31 (4.53)	-24.89 (3.27)	291 (133.84)	0.110 (0.035)	4.37 (1.53)
<i>p</i> -Value of Wald Test of Equality		0.000	0.805	0.026	0.098	0.051
Nationality	German	95.33 (4.19)	-26.20 (3.03)	312 (118.11)	0.118 (0.032)	2.80 (1.51)
	Foreign	94.52 (5.26)	-20.64 (3.69)	46 (58.59)	0.029 (0.035)	4.83 (1.65)
<i>p</i> -Value of Wald Test of Equality		0.829	0.048	0.000	0.000	0.048
Age	Age 25-34	105.07 (4.21)	-21.47 (3.66)	183 (91.28)	0.087 (0.032)	3.11 (1.47)
	Age 35-44	84.83 (4.64)	-28.68 (3.25)	236 (96.94)	0.088 (0.032)	3.57 (1.48)
	Age 45+	93.66 (5.31)	-21.00 (4.07)	228 (88.52)	0.113 (0.034)	3.51 (1.53)
<i>p</i> -Value of Wald Test of Equality		0.000	0.018	0.006	0.346	0.059
Education	No voc. degree and missing	93.04 (4.45)	-22.45 (3.57)	49 (52.05)	0.043 (0.031)	4.12 (1.47)
	Vocational training degree	92.77 (4.42)	-25.03 (3.08)	311 (130.21)	0.120 (0.034)	3.32 (1.56)
	University degree	125.15 (5.84)	-34.09 (6.48)	350 (213.10)	0.099 (0.045)	1.85 (1.35)
<i>p</i> -Value of Wald Test of Equality		0.000	0.129	0.005	0.000	0.005
Occupation	Elementary occupations	93.46 (5.37)	-20.58 (4.59)	60 (54.53)	0.022 (0.034)	5.53 (1.50)
	Agricultural occ. and missing	97.80 (8.24)	-45.47 (13.15)	47 (56.98)	0.077 (0.039)	4.09 (1.62)
	Plant/machine operator	89.99 (4.67)	-23.17 (5.01)	54 (56.86)	0.078 (0.031)	5.06 (1.64)
	Craft and related occupations	90.58 (4.61)	-30.85 (4.02)	64 (59.14)	0.087 (0.036)	3.20 (1.93)
	Service worker	99.54 (6.57)	-11.59 (5.33)	161 (90.64)	0.073 (0.033)	6.80 (1.55)
	Clerk	89.61 (4.87)	-28.53 (4.06)	354 (151.83)	0.111 (0.039)	1.63 (1.69)
	Techn., assoc. professional	99.84 (4.28)	-22.16 (4.19)	432 (170.25)	0.166 (0.038)	1.22 (1.66)
	Professional	106.97 (5.72)	-29.37 (5.88)	207 (167.75)	0.091 (0.040)	0.33 (1.43)
	Senior official, manager	139.72 (11.63)	-33.95 (13.24)	427 (248.20)	0.166 (0.155)	8.60 (2.04)
<i>p</i> -Value of Wald Test of Equality		0.000	0.007	0.075	0.000	0.000

Notes: Unemployment and employment durations are measured in days, earnings per calendar day in € (real values, reference year 2000). P50 refers to the difference between treatment and nontreatment outcomes at the median of the cumulative distribution functions. Calculations are based on simulated outcomes with and without training for a subset of the originally treated persons. Bootstrapped standard errors based on 250 replications in parentheses.

# Online Appendix to “Quick Job Entry or Long-Term Human Capital Development? The Dynamic Effects of Alternative Training Schemes”

ADERONKE OSIKOMINU

(University of Zurich, Albert-Ludwigs-University Freiburg, CESifo and IZA)

## A Detailed Identification Analysis

In this section, we lay out the details of the identification analysis of the model described in section 3.2. Our goal is to identify the single components of the hazard rates  $\theta_{kl}(\cdot)$  defined in equation (3) and the joint distribution of the unobserved heterogeneity terms, which then yields the joint distribution of the survival times. The identification analysis is somewhat nonstandard as different parts of our model require different identifying assumptions. We start with the probability of survival in the initial state and then move on to the higher order transitions.

Let  $T_{Ol} \in \mathbb{R}_+$  be the survival time in the initial state  $O$  until exit to state  $l \in \mathcal{Z}^a(O) \equiv \{S(O), L(O), E(O)\}$ , where  $S$  denotes short-term training,  $L$  long-term training and  $E$  employment. The state in parentheses, i.e.  $(O)$ , indicates dependence of current transition rates on the past trajectory. Let  $x(t)$  denote the covariate process just before time  $t$ ,  $\bar{x}(t) \equiv \{x(r), 0 \leq r < t\}$  the covariate path on  $[0, t)$ , and  $X$  its support. Integrating the hazard rate in equation (3) over  $t$  we obtain the integrated hazard:

$$(1) \quad \Theta_{Ol}(t|\bar{x}(t), v_{Ol}) \equiv \left( \int_0^t \lambda_{Ol}(r, x(r)) dr \right) v_{Ol} = \Lambda_{Ol}(t, \bar{x}(t)) v_{Ol},$$

where the function  $\Lambda_{Ol}(t, \bar{x}(t)) : (\mathbb{R}_+ \times X) \rightarrow \mathbb{R}_+$  is the integrated structural hazard that summarizes the dependence on observed factors and  $v_{Ol}$  is a non-negative unobserved heterogeneity term with support  $V_{Ol}$ . Conditional on the observed covariates the probability of survival in the initial state is:

$$\begin{aligned} & \Pr(\cap_{l \in \mathcal{Z}^a(O)} T_{Ol} > t | \bar{x}(t)) \\ &= \int_{V_{OS(O)}} \int_{V_{OL(O)}} \int_{V_{OE(O)}} \exp \left[ - \sum_l \Lambda_l(t, \bar{x}(t)) \nu_{Ol} \right] dG_O^1(\nu_{OS(O)}, \nu_{OL(O)}, \nu_{OE(O)}) \\ &= \mathcal{L}_{G_O^1} [\Lambda_{OS(O)}(t, \bar{x}(t)), \Lambda_{OL(O)}(t, \bar{x}(t)), \Lambda_{OE(O)}(t, \bar{x}(t))] , \end{aligned}$$

where  $G_O^1$  is the cumulative distribution function of  $(v_{OS(O)}, v_{OL(O)}, v_{OE(O)})$  and  $\mathcal{L}_{G_O^1}(s_1, s_2, s_3)$  its trivariate Laplace transform. With competing destination states, the joint and marginal survival probabilities are not observed but the sub-survival probabilities

$$Q_{Ol}(t|\bar{x}(t)) \equiv \Pr [T_{Ol} > t, \cap_{m \in \mathcal{Z}^a(O) \setminus l} T_{Om} > T_{Ol} | \bar{x}(t)]$$

and the sub-survival densities

$$\begin{aligned} Q'_{Ol}(t|\bar{x}(t)) &\equiv \partial Q_{Ol}(t|\bar{x}(t))/\partial t \\ (2) \quad &= \lambda_{Ol}(t, x(t)) D_{Ol} \mathcal{L}_{G_O^1}(\Lambda_{OS(O)}(t, \bar{x}(t)), \Lambda_{OL(O)}(t, \bar{x}(t)), \Lambda_{OE(O)}(t, \bar{x}(t))), \end{aligned}$$

where  $D_i \mathcal{L}_{G_O^1}(s_1, s_2, s_3) \equiv \partial \mathcal{L}_{G_O^1}(s_1, s_2, s_3) / \partial s_i$ ,  $i = 1, 2, 3$ . We have that the probability of survival in the initial state  $\Pr(\cap_{l \in \mathcal{Z}^a(O)} T_{Ol} > t) = \sum_l Q_{Ol}(t)$ , Tsiatis (1975, Theorem 1). Identification of the subdensities (2) and  $G_O^1$  is discussed in Abbring and van den Berg (2003b). In the model with employment as the second outcome we use repeated cycles per individual. In this case, identification does not require variation with observed covariates (Abbring and van den Berg 2003b, Proposition 3).

In the model with earnings as the second outcome we have only one cycle per individual. Therefore, we impose more structure on the variation of the hazard rates with observed covariates. We model the hazard rates as proportional in the three components a flexible nonnegative function of elapsed duration,  $\lambda_{Ol}^0(t)$ , of observed covariates,  $\phi_{Ol}(x(t))$ , and unobserved heterogeneity,  $v_{Ol}$ , i.e.:  $\lambda_{Ol}(t, \bar{x}(t)) = \lambda_{Ol}^0(t) \phi_{Ol}(x(t)) v_{Ol}$ . Further, we assume that  $\bar{x}(t)$  includes three continuous covariates that are independent of  $v_{Ol}$ ,  $l \in \mathcal{Z}^a(O)$ , and that  $(\phi_{OS(O)}(x(t)), \phi_{OL(O)}(x(t)), \phi_{OE(O)}(x(t)))$  varies on a nonempty open set in  $\mathbb{R}_+^3$ . In the model with earnings, the probability of survival in the initial state is identified applying Abbring and van den Berg (2003b, Proposition 2) with one exception. As  $x(t)$  includes time-varying covariates we do not need finiteness of the mean of the unobservables, see Heckman and Taber (1994, Theorem 4) replicating a proof in Honoré (1990).

In contrast to (2), the subdensities for the second and higher order transitions depend on lagged duration in unemployment,  $\tau$ . The conditional sub-survival densities for transitions from the second state in a cycle  $k \in \{S(O), L(O), E(O)\}$  equal:

$$\begin{aligned} Q'_{kl}(t|\tau, \bar{x}(t); v_{OS(O)}, v_{OL(O)}, v_{OE(O)}) \\ = \lambda_{kl}(t, \tau, x(t)) \times \\ D_{kl} \mathcal{L}_{G_k^2}(\{\Lambda_{kl}(t, \tau, \bar{x}(t)) : l \in \mathcal{Z}^a(k)\}), \end{aligned}$$

where  $G_k^2 = G_k^2(\{v_{kl} : l \in \mathcal{Z}^a(k)\} | v_{OS(O)}, v_{OL(O)}, v_{OE(O)})$ . As regards the transition rates from  $L(O)$  and  $E(O)$  we can directly apply the result from Brinch (2007) to identify  $G_k^2(v_{kl} | v_{OS(O)}, v_{OL(O)}, v_{OE(O)})$  and  $\Lambda_{kl}(\cdot)$ . Brinch shows that in a mixed hazard model with univariate unobserved heterogeneity and time-varying covariates the survival distribution is nonparametrically identified without functional form restrictions beyond separability of the hazard rate in the unobserved heterogeneity term. Further, we model the hazard rate from  $S(O)$  to  $L(S)$  as  $\lambda_{OL(O)}(\tau + t, x(t))\phi_{S(O)L(S)}(x(t))v_{OL(O)}$ , where  $\phi_{S(O)L(S)}(\cdot)$  is some flexible nonnegative function.<sup>1</sup>  $\lambda_{OL(O)}(\cdot)$  and the distribution of  $v_{OL(O)}$  have been identified previously.  $\phi_{S(O)L(S)}(x(t))$  can be identified following Abbring and van den Berg (2003a, Proposition 3). With this restriction, we can apply Brinch's strategy to identify the transition rate from  $S(O)$  to  $E(S)$ , too. In practice, we model the transition rates from treated unemployment,  $k \in \{S(O), L(O)\}$ , to employment,  $E(k)$ , as  $\theta_{kE(k)}(t, \tau, \bar{x}(t), v_{kE(k)}) = \lambda_{OE(O)}(\tau + t, x(t))\phi_{kE(k)}(\tau, x(t))v_{kE(k)}$ . In this way, we align treated and controls with respect to their unemployment experience.

In our alternative model with earnings as the second outcome, we do not have covariates that vary with the level of earnings. We can apply the identification results of Honoré (1993, Theorem 3) or, alternatively, Horny and Picchio (2010) to the subdensity of earnings. Here, we need additional proportionality restrictions, i.e.  $\theta_{kE(k)}(t, \tau, x, v_{kE(k)}) = \lambda_{kE(k)}^0(t)\kappa_{kE(k)}(\tau)\phi_{kE(k)}(x)v_{kE(k)}$ , with  $\lambda_{kE(k)}^0(\cdot)$ ,  $\kappa_{kE(k)}(\cdot)$ ,  $\phi_{kE(k)}(\cdot)$  some flexible nonnegative functions, and additional assumptions on the variation with observed covariates, analogous to those on the hazard rates from the initial state in the model with earnings.

We proceed analogously with the remaining transitions in a cycle. Once we know all the integrated structural hazards,  $\Lambda_{kl}(\cdot)$ , we can trace out the joint distribution of the unobservables,  $G(\mathbf{v})$ , by varying the observed arguments (i.e.  $t, \tau, \bar{x}(t)$ ) of the hazard rates on a nonempty open set in  $\mathbb{R}_+^p$ , where  $p$  is the dimension of  $\mathbf{v}$ , Abbring and van den Berg (2003b, Proposition 1). Due to the analyticity of the Laplace transform we only need locally independent variation of the hazard rates with observables in order to identify the joint distribution of the unobservables on the entire domain. This requirement is substantially weaker than corresponding assumptions on exclusion restrictions in discrete choice models.

---

<sup>1</sup>In principle, we could allow  $\phi_{S(O)L(S)}(\cdot)$  to depend on  $\tau$  and  $t$  in addition to  $x(t)$ .



## B Institutional Background

### B.1 Unemployment, Job Search, and Job Training

Since the late 1990s public employment services in Germany increasingly emphasize activation and individual responsibilities of job-seekers.<sup>2</sup> People who seek support from public employment services are required to undertake significant independent efforts towards finding a job. In particular, they have to attend counseling interviews, provide proof of their independent job search, and participate in active labor market programs. On legal grounds, a caseworker has to provide detailed instructions to each particular job-seeker on what it takes for him/her to be actively looking for a job (Schweiger, 2002). Typically, this includes a specification of the number of written applications a job-seeker has to send out within a given period of time. Such explicitness is necessary in order to create a legal basis for monitoring compliance with search requirements. Job-seekers who fail to cooperate with the local employment agency may receive punitive benefit sanctions and be denied further support.

While, during the period under study (i.e. 1999-2004), the regulation of public employment services is very specific on the importance of active job search, the link between job training, individual needs of job-seekers, and demands of the labor market is very vague. Efficiency considerations play a limited role as job training also serves social purposes. Rather, public sponsored training aims at improving human capital as an end in itself.<sup>3</sup> Furthermore, training programs, in particular short-term training schemes, are used as an instrument to monitor the availability and willingness to work.

The allocation of training schemes, as almost all active labor market programs, is subject to the discretionary power of the caseworker at the local employment agency. Registered job-seekers have no legal entitlement to receive any particular program (Bieback, 1997). The caseworker has to judge participation in training as necessary in order to improve the employment prospects of the job-seeker. The job-seeker and the caseworker meet repeatedly during the unemployment spell to discuss further search strategies and program participation. The unemployed may also just receive a written notification to participate in a program. Training programs start at many

---

<sup>2</sup>For a summary of recent institutional developments of German labor market policy see Seeleib-Kaiser and Fleckenstein (2007), for a detailed overview see Eichhorst et al. (2007).

<sup>3</sup>In this context, experts at local employment agencies talk of ‘stockpiling of human capital’ (“Bildung auf Vorrat”, Schneider et al., 2006, p. 77).

different times during the year.<sup>4</sup> Registered job-seekers may participate in a program at any point in time during their unemployment spell. Participation is mandatory and noncompliance may entail a temporary suspension of benefits. However, job placement has priority over program participation. The unemployed are therefore encouraged to continue job search at any time, even while participating in a training program.

The caseworker's assignment decision is to a large extent driven by the supply of courses. In an informal procedure, the employment agencies and the training providers negotiate in advance the quantities and contents of training courses to be offered during the next calendar year. A high utilization of planned training capacities is an important administrative target at the local employment agencies.<sup>5</sup> Blaschke and Plath (2000) report that also private indicators of the caseworker like the composition of a group of participants in a particular course or his/her assessment of the motivation of the unemployed play an important role. However, caseworkers face a very high caseload which makes a careful targeting of job-seekers to particular training programs or informal agreements between job-seekers and caseworkers impossible.<sup>6</sup> The assignment to training programs occurs in general at short notice. On the one hand short notice periods and belated assignments (i.e. after the official starting date of a program) are used as a work test. On the other hand they are convenient for caseworkers who want to fill upcoming training slots as well as maintain incentives for independent job search at a high level.<sup>7</sup>

---

<sup>4</sup>Some training schemes also do not have fixed start and end dates in the sense that a group of participants starts a course with a fixed scheduled duration together at some given date. Rather, the schemes continue over an indefinite period of time and new participants enrol for a varying scheduled duration determined by the caseworker as soon as previous participants exit the program (Blaschke and Plath, 2000).

<sup>5</sup>Experts at local labor agencies refer to the process of training assignment commonly as 'filling/feeding of training slots' ("Maßnahmebeschickung", "Maßnahmefüllung", Schneider et al., 2006, pp. 77, 98). Further evidence in Schneider et al. (2006) documents that planned and actually used training capacities nearly coincide.

<sup>6</sup>In 2001 and 2002, the average caseload in terms of registered unemployed is about 400 unemployed per caseworker (Bundesanstalt für Arbeit, 2002, p. 87). Mosley et al. (2002, p. 4) report a caseload between 267 and 580 unemployed per caseworker in 2000 in West Germany. In addition, caseworkers counsel people who register as job-seekers but are not unemployed as well as people who do not register in the end.

<sup>7</sup>While the assignment of short-term training programs remained unchanged throughout our analysis period, since 2003, long-term training schemes are assigned according to a somewhat different regime. The reform affects about 10% of the long-term training programs in our sample. Since the reform, candidates for a long-term training program obtain a voucher that details a certain training field. The job-seeker then redeems the voucher by choosing a suitable course from a pool of certified courses. The voucher reform intended to make the allocation of training programs more targeted and better tailored towards the needs of the unemployed. However, its original goals were impaired by important difficulties in the implementation of the new system. In particular, it turned out that training providers tended to collect vouchers until a sufficient critical

## B.2 Aggregate Figures on Participation and Expenditures

Active labor market policy in Germany comprises a variety of programs ranging from subsidized employment on the first labor market, job creation schemes on the second labor market to job training programs that adjust and enhance the qualifications of participants. Employment subsidies aim at promoting either dependent employment or business start-ups. Job creation schemes provide employment opportunities in non-profit organizations for long-term and difficult-to-place unemployed. In addition to these large scale programs, there also exist programs targeted to particular groups as e.g. young or disabled persons. Active labor market policy is complemented by placement and advisory services, that are increasingly contracted out to private providers.

Long-term training schemes have traditionally been the most important field of active labor market policy in Germany. Since 1998 there have been several reforms leading to a focus on measures considered particularly effective in activating the unemployed in the short run and in preventing long-term unemployment. Thus, allocation of resources shifted away from human capital intensive long-term training to job-search oriented short-term training. Figure B1 shows a decline in entries in the long-term training programs – in total Germany (panel (a)) as well as in the Western Länder (panel (b)) – whereas participation in short-term training increases over time.

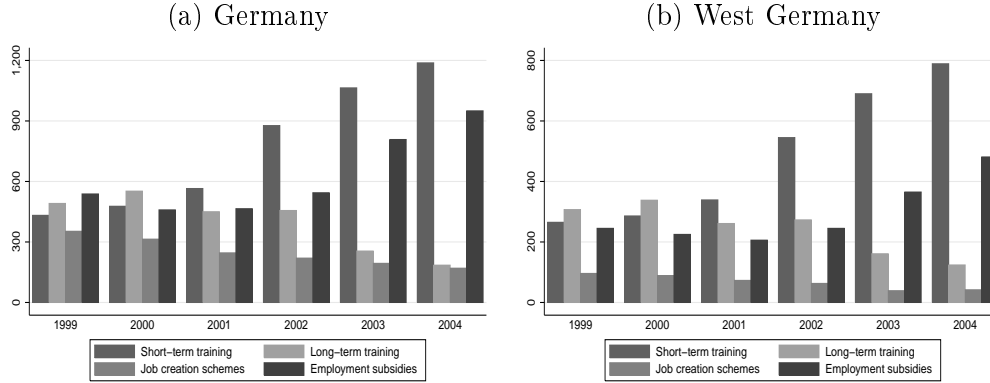
During the first half of the 2000s, real public spending on active and passive labor market policy is on average €70 billion per year, about 3.1 to 3.4% of GDP. About €20 billion per year are spent on active labor market policies, a quarter of that amount go to training (Bundesagentur für Arbeit, 2001a, 2003a, 2005a, table II.E.6).<sup>8</sup> Table B1 provides a breakdown of expenditures on training and unemployment insurance payments per person-month as well as average program durations. It shows that the average monthly training costs per participant are slightly lower for short-term training courses (560 Euros on average) than for traditional long-term training schemes (650 Euros on average). Most striking is the great difference in

---

number of participants was reached or they shortly canceled scheduled courses if there were too few participants (Kühnlein and Klein, 2003; Schneider et al., 2006). According to surveys conducted by the Federal Institute for Vocational Education and Training (BIBB) among training providers, in 2003 and 2004, 50% and 60%, respectively, of the offered courses had to be canceled shortly because of an insufficient number of participants (Paulsen et al., 2006). Therefore, potential participants faced a high uncertainty regarding the start of a long-term training program.

<sup>8</sup>According to OECD (2006, ch. 7), public spending on active labor market policies per unemployed relative to per capita GDP in the years 2000 and 2001 is around 33% in Germany, which is at the mean across OECD countries.

Figure B1: Entries into Selected Active Labor Market Programs (in 1000)



Source: Bundesanstalt für Arbeit (2001a, 2003a), Bundesagentur für Arbeit (2005a); own calculations.

average duration of the courses, that is displayed in the lower panel of table B1. While short-term training courses last on average one month, the mean duration of long-term training programs lies between eight and ten months. Under the budgetary pressure caused by a persistently high unemployment rate and in light of these large differences in costs, the share of short-term training measures drastically increased in 2002 and, in 2003 and 2004, this rise continues at the expense of the traditional longer-term measures. Of course, the higher training costs of the latter would be justified if they were associated with correspondingly higher gains.

## C Further Details on the Data Used

This study uses data from the Integrated Employment Biographies (IEB) Version 4.02 which are compiled by the Research Data Center of the German Federal Employment Agency.<sup>9</sup> The IEB is an integrated administrative database combining records out of four different registers: the Employment History (*Beschäftigten-Historik*), the Benefit Recipient History (*Leistungsempfänger-Historik*), the Job-Seeker Database (*Bewerberangebot*), and the Program Participation History (*Maßnahme-Teilnehmer-Gesamtdatenbank*). The basic population consists of all individuals who, during the period 1990 to mid-2005, have either held a job sub-

<sup>9</sup>For descriptions of the data see Jacobebbinghaus and Seth (2007) and <http://fdz.iab.de/en.aspx>. The data are subject to confidentiality regulations. A weakly anonymous version can be used on-site and via controlled remote access, a factually anonymous version is available as scientific use file, in which regional information, nationality and industry are provided at a more aggregated level. The analysis in this paper is based on a weakly anonymous version.

Table B1: Average Monthly Expenditures and Program Durations

	1999	2000	2001	2002	2003	2004
(a) Average monthly expenditures (in €)						
Short-term training	602	580	570	658	538	421
Long-term training	1,570	1,627	1,668	1,686	1,555	1,574
– subsistence allowance	1,093	1,152	1,178	1,188	1,156	1,150
– training costs	629	640	664	681	631	627
Unemployment benefits	1,132	1,160	1,189	1,185	1,261	1,313
Unemployment assistance	869	753	721	727	691	713
(b) Average program duration (in months)						
Short-term training	1.1	1.2	1.1	0.9	1.0	0.9
Long-term training	8.4	8.2	9.3	9.1	10.5	10.7

Notes: Panel (a) contains the average monthly expenditures per participant/benefit recipient, panel (b) the average program duration in months. Expenditures on subsistence allowance, unemployment benefits and unemployment assistance include social security contributions. Source: Bundesanstalt für Arbeit (2000, 2001b, 2002, 2003b), Bundesagentur für Arbeit (2004, 2005b, 2005c); own calculations.

ject to social security contributions or have been registered as a benefit recipient, job-seeker, or program participant at an employment agency.

The Employment History and the Benefit Recipient History cover the time period January 1990 to December 2004 (employment) and June 2005 (benefits), respectively. The Employment History records employment spells subject to social security contributions as well as marginal part-time employment partly exempted from social security contributions. Approximately 80% of all employed persons in West Germany hold jobs that pay social security contributions (Bender et al. 2000, p. 651). We calculate the individual entitlement periods to unemployment benefits using the Employment History and the Benefit Recipient History and a summary of the different regulations provided by Plaßmann (2002). The Program Participation History records active labor market programs starting in the period 2000 to mid-2005 and partly also those taking place in earlier years. Comparing the entries into different programs in 1999 with the figures for later years shows that information on programs starting in 1999 seems to be already complete for most active labor market programs. We use information on the planned end dates of training programs from this database. About 7% of the long-term training spells and 25% of the short-term training spells have missing or implausible planned end dates. In these cases, we impute a planned end date based on the median planned duration within disaggregated, five-digit program codes. The Job-Seeker Database completely records job search spells starting between January 2000 and June 2005

and partly also those beginning before 2000 if the person in question keeps the same client number throughout.

The data are constructed as an event history data set with start and end dates measured on a daily basis. An important feature of the data set is that it contains multiple parallel spells in the case of concurrent events. This allows us e.g. to distinguish regular employment from subsidized employment in the context of active labor market policy. Moreover, the IEB comprises a large set of variables that provide detailed information on socioeconomic, occupational and job characteristics, as well as of job search and contents of active labor market programs.

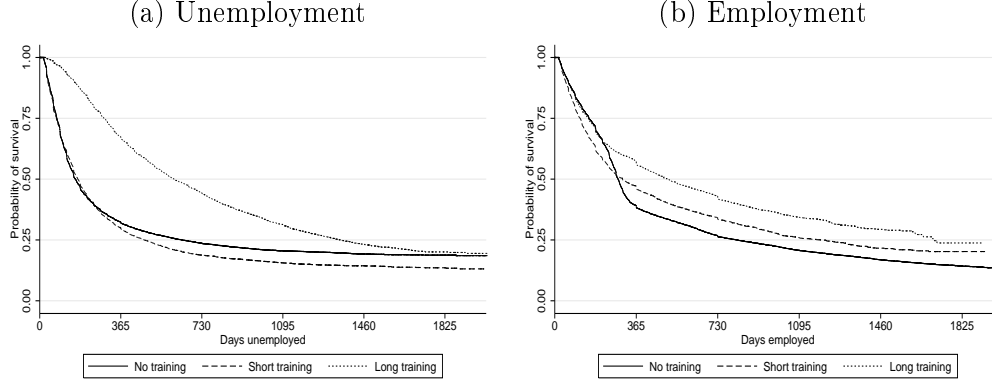
## D Detailed Descriptive Analysis

### D.1 Descriptive Evidence on Outcomes with and without Training

Figure D1 shows Kaplan-Meier estimates of the probability to survive in unemployment and employment beyond the number of days in unemployment and employment, respectively, given on the abscissa separately by treatment status. These estimates do not correct for selective training participation or employment take up. Treatment status during unemployment is a time-varying variable, which means that future participants contribute to the nontreated survivor function until they start training. In panel (a) of figure D1, we see that the curve for the treated with long-term training lies above the one for the non-treated and that the area between the two lines has the shape of a lens. This suggests that, initially, people enrolling into long-term training leave unemployment at a slower rate than nontreated individuals, which is reflected in a widening of the vertical distance between the two curves. As the share of treated who have completed their training program increases with elapsed unemployment duration the vertical distance between the dotted and the solid curve eventually narrows again, implying that treated persons now leave unemployment at a faster rate than non-treated. For short-term training there is no evidence of such a lock-in effect during program participation. Even at short elapsed unemployment durations the survivor functions of participants in short-term training and nontreated persons are very close. At higher elapsed durations, the fraction of treated with short-term training still unemployed is clearly smaller than that of nontreated or that of treated with long-term training.

Panel (b) of figure D1 shows the survivor functions in employment by treatment

Figure D1: Probability of Survival in Unemployment and Employment



Notes: Kaplan-Meier (1958) estimates of the probability to remain unemployed and employed, respectively, beyond the number of days given on the abscissae. During unemployment, treatment status varies with time, i.e. future participants contribute to the nontreated survivor function until they start training. Calculations are based on 87,250 unemployment spells and 56,758 employment spells, respectively.

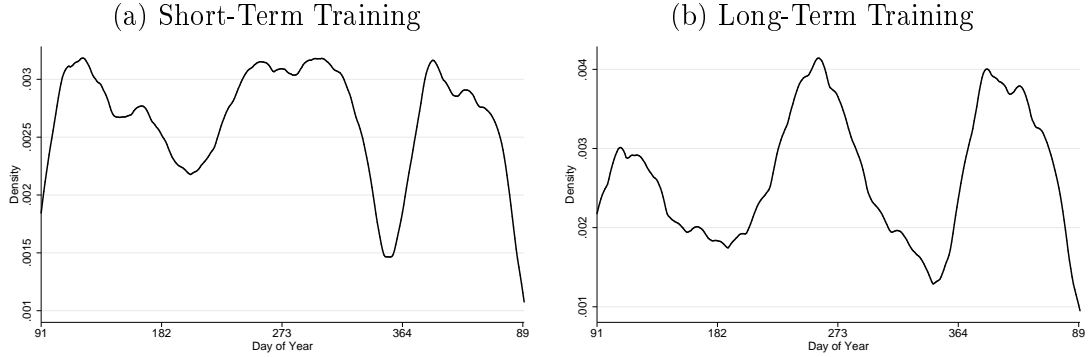
status in the preceding unemployment spell. At short elapsed durations, individuals treated with short-term training exit employment at a faster rate than nontreated or former participants in long-term training. The latter have the most stable employment relationships. In fact, the vertical distance to the survivor functions of persons treated with short-term training and nontreated first increases and then remains largely stable at higher elapsed durations.

## D.2 Variation of Training Starts with Time

In the timing-of-events approach we trace out treatment effects by comparing individuals with the same elapsed unemployment duration but different training starts. Somebody who has not yet enrolled in training serves as a control for somebody who has already started training. The framework relies thus strongly on variation of training starts with time. The institutional setup places no restrictions on when programs should start in the calendar year. It also places no restrictions on the point in time during unemployment. Figure D2 shows the density of training starts over the calendar year. We notice a short dip at Christmas and a reduced number of program starts in June/July but in general training programs start on all days throughout the year. Figure D3 displays the Kaplan-Meier estimates of the probability to start a program after some given time in open unemployment. The unit of measurement is a calendar day. It can be seen that the hazard rate into short-term training varies between 0.015 and 0.056% and that of long-term training lies between

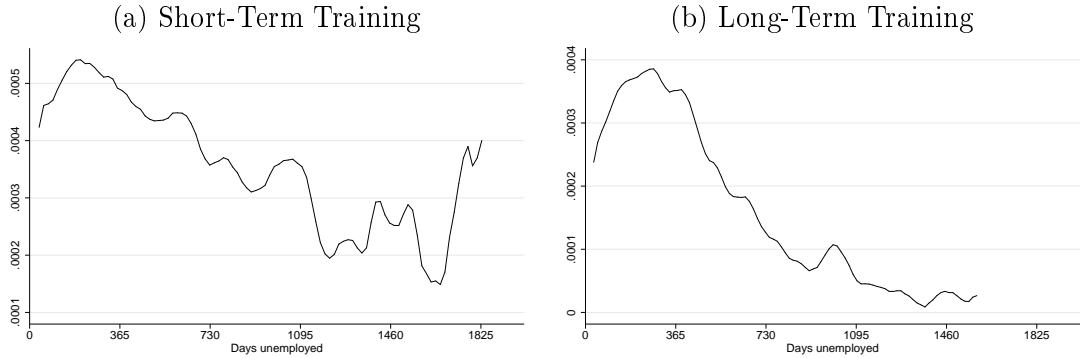
0.003 and 0.037%.

Figure D2: Density of Program Starts throughout the Year



Notes: The bandwidth for the kernel smooth of the density is 10 days. Calculations are based on 8,279 and 5,580 starts of short-term and long-term training, respectively.

Figure D3: Hazard Rate into Training Programs



Notes: Kaplan-Meier (1958) estimates of the probability to start training within a short period of time conditional on unemployment lasting until the number of days given on the abscissae. The bandwidth for the kernel smooth of the hazard rate is 30 days. Boundary estimates within 30 days of the endpoints are omitted. Calculations are based on 87,250 unemployment spells.

Table D1 presents descriptive statistics on the elapsed unemployment duration at training start and the planned durations of short-term and long-term training, respectively. The majority of training starts occurs within the first four months of unemployment, i.e. the median waiting time is 140 for either training program. The distribution of training starts is highly dispersed and right skewed with a mean waiting time of around 200 days for both short-term and long-term training and standard deviations of 261 days for short-term training and 193 days for long-term training. Table D1 also shows the distribution of planned program durations. The planned program duration is assigned at the start of the program and may differ from ex post realized duration. The median duration of short-term training is four



weeks while for long-term training it is 6.5 months.

Table D1: Distribution of Training Starts and Planned Program Duration

	Mean (SD)	P25	P50	P75
Short-Term Training				
Waiting time	228.72 (260.90)	52	140	304
Program duration	32.89 (26.46)	12	26	54
Long-Term Training				
Waiting time	197.18 (192.49)	60	142	277
Program duration	312.44 (250.54)	176	201	362

Notes: Calculations are based on 8,279 starts of short-term training and 5,580 starts of long-term training. Durations are measured in days. Program duration refers to the planned program duration assigned at the start of training. It may differ from the ex post realized training duration. SD denotes the standard deviation. P25, P50, P75 refer to the 25th, 50th, 75th percentile of the cumulative distribution function.

### D.3 Descriptive Statistics for Key Variables

Table D2: Means and Standard Deviations

Variable Description	No Training	Short Training	Long Training
Demographics			
Female	0.398 (0.489)	0.432 (0.495)	0.467 (0.499)
Foreign	0.388 (0.487)	0.358 (0.479)	0.312 (0.463)
Ethnic German	0.025 (0.157)	0.021 (0.142)	0.033 (0.178)
Has children	0.343 (0.475)	0.351 (0.477)	0.367 (0.482)
Age	37.609 (8.007)	39.455 (8.133)	39.061 (7.547)
Family status not available (na)	0.121 (0.326)	0.105 (0.307)	0.096 (0.295)
Single	0.348 (0.476)	0.368 (0.482)	0.349 (0.477)
Cohabiting	0.035 (0.184)	0.043 (0.202)	0.041 (0.199)
Lone parent	0.039 (0.193)	0.058 (0.234)	0.069 (0.253)
Married	0.457 (0.498)	0.426 (0.495)	0.445 (0.497)
Entitlement to Unemployment Benefits			
Has claim	0.902 (0.297)	0.912 (0.284)	0.937 (0.243)
Total claim (in 30 days)	11.518 (4.637)	12.028 (4.906)	11.569 (4.402)
Formal Education and Occupation in Last Job			
School degree na	0.015 (0.123)	0.012 (0.109)	0.027 (0.163)
No school degree	0.105 (0.306)	0.120 (0.325)	0.078 (0.267)
Haupt-/Realschule (grade 9-10)	0.717 (0.451)	0.712 (0.453)	0.692 (0.462)
(Fach-)Abitur (grade 12-13)	0.163 (0.369)	0.156 (0.363)	0.204 (0.403)
Vocational degree na	0.015 (0.120)	0.010 (0.102)	0.024 (0.154)
No vocational degree	0.335 (0.472)	0.381 (0.486)	0.336 (0.473)
Apprenticeship degree	0.588 (0.492)	0.551 (0.497)	0.574 (0.495)

<continued on next page>

Table D2: Means and Standard Deviations &lt;continued&gt;

Variable Description	No Training	Short Training	Long Training
University degree	0.062 (0.242)	0.057 (0.233)	0.065 (0.247)
Occupation na	0.005 (0.068)	0.003 (0.054)	0.005 (0.068)
Elementary occupations	0.149 (0.356)	0.166 (0.373)	0.142 (0.349)
Agricultural occupations	0.028 (0.164)	0.019 (0.137)	0.016 (0.126)
Plant/machine operator	0.116 (0.320)	0.113 (0.316)	0.117 (0.321)
Craft and related workers	0.228 (0.420)	0.198 (0.399)	0.154 (0.361)
Service/sales worker	0.149 (0.356)	0.149 (0.356)	0.133 (0.339)
Clerk	0.141 (0.348)	0.178 (0.383)	0.234 (0.423)
Techn., assoc. professional	0.096 (0.294)	0.100 (0.299)	0.112 (0.316)
Professional	0.073 (0.261)	0.059 (0.235)	0.075 (0.263)
Senior official, manager	0.017 (0.128)	0.015 (0.121)	0.013 (0.114)
Caseworker's Assessment			
Skill level in target job na	0.062 (0.242)	0.033 (0.180)	0.075 (0.264)
Unskilled in target job	0.398 (0.489)	0.452 (0.498)	0.398 (0.490)
Medium skills in target job	0.467 (0.499)	0.444 (0.497)	0.454 (0.498)
High skills in target job	0.073 (0.260)	0.071 (0.256)	0.072 (0.258)
Experienced in target job	0.894 (0.308)	0.914 (0.280)	0.874 (0.332)
Health status na	0.051 (0.221)	0.017 (0.129)	0.035 (0.185)
No health constraints	0.828 (0.377)	0.830 (0.375)	0.828 (0.377)
Health constr. without consequ.	0.059 (0.236)	0.074 (0.261)	0.068 (0.252)
Health constraints with consequences for placement	0.061 (0.240)	0.079 (0.270)	0.068 (0.252)
Observations	37,529	3,827	4,103

Notes: All variables are dummies except age and total claim. All variables refer to the start of the first unemployment spell. Standard deviations are in parentheses. Long-term training includes people who first participate in short-term training.

## E List of Variables Used in the Estimation

Table E1: Variable Definitions

Name	Definition
bhazXX	time-varying dummy equal to one if elapsed duration (in days) is greater than XX days and smaller than or equal the number of days referring to the next time interval dummy; the last interval is open ended
f_bhazXX, fo_bhazXX, c_bhazXX	baseline hazard dummy interacted with female, foreign, indicator for repeated cycle, respectively
female	dummy equal to one if female
foreign	dummy equal to one if citizenship is not German

<continued on next page>

Table E1: Variable Definitions &lt;continued&gt;

Name	Definition
ethnicgerman	dummy equal to one if ethnic German, i.e. returned settler from former German settlements
agegroup	age in 6 categories: 1 25-29 years, 2 30-34 years, 3 35-39 years, 4 40-44 years, 5 45-49 years, 6 50 or older; time-varying
age, agesq	age divided by ten, age squared divided by 100, centered around 38 years, time-varying
f_age, f_agesq	age, agesq interacted with female
education	1 no information available, 2 no degree, 3 vocational training degree, 4 university or technical college degree
schooling	1 no information available, 2 no schooling degree, 3 Hauptschulabschluss or Mittlere Reife/Fachoberschule (degrees reached after completion of the 9th or 10th grade), 4 Fachhochschulreife or Abitur/Hochschulreife (degrees reached after completion of the 12th or 13th grade)
health	when unemployed with four categories: 1 no information available, 2 no health constraints mentioned, 3 health constraints, but considered without consequences for placement, 4 health constraints considered to have an impact on placement; when employed dummy equal to one if person had health constraints affecting placement within the last two months of job search before start of employment spell
disabled	dummy equal to one if disabled
family	1 no information available; 2 living alone; 3 not married, but living together with at least one other person; 4 lone parent; 5 married
kids	dummy equal to one if person has children
youngchild	dummy equal to one if person has children younger than 10 years
occupation	occupation (of last employment) in 10 categories: 1 no information available, 2 elementary occupations, 3 agricultural occupations, 4 plant/machine operator, 5 craft and related workers, 6 service/sales workers, 7 clerk, 8 technicians and associate professionals, 9 professionals, 10 senior officials and managers
whitecollar	dummy equal to one if white-collar job
bluecollar	dummy equal to one if blue-collar job
parttime	dummy equal to one if weekly hours worked less than full-time
lwquart1-lwquart4	dummy variables indicating the quartile of last salary
lnlwage, lnlwagesq	log of last real salary and its square if salary is below social security threshold, else zero
industry	industry (of last employment) in 7 categories: 1 no information available; 2 agriculture, forestry, fishing; 3 manufacturing; 4 construction; 5 trade and transport; 6 financial, renting and business; 7 other services
seasonwork	dummy equal to one if industry (of last employment) characterized by seasonal work

&lt;continued on next page&gt;

Table E1: Variable Definitions &lt;continued&gt;

Name	Definition
endlastjob	1 other reason and no information available, 2 termination of last employment by employer, 3 by employee, 4 fixed term contract
prevtrans	dummy equal to one if received some kind of unemployment insurance benefits in the three years preceding the current unemployment spell
earlycontact	dummy equal to one if already registered as job-seeker up to three months before beginning of current unemployment spell
taredu	1 no information available, 2 caseworker considers job-seeker as low skilled in target profession, 3 considered as medium skilled, 4 considered as high skilled
tarexp	dummy equal to one if caseworker considers job-seeker to have professional experience in target profession
seekpt	dummy equal to one if seeking only parttime job
sick	time-varying dummy equal to one if temporarily sick (unemployment spells only)
eversick	time-varying (time-constant in employment spells) dummy equal to one if experienced a period of sickness during the current or past unemployment spell
offben	time-varying dummy equal to one if temporarily off unemployment transfers because of sanctions (unemployment spells only)
eversanction	time-varying (time-constant in employment spells) dummy equal to one if received punitive benefit sanction during the current unemployment or past unemployment spell
hasclaim	time-varying dummy equal to one if entitled to unemployment benefits and claim not yet expired (unemployment spells only)
totclaim, totclaimsq	time-varying variable (time-constant in employment spells), (square of) total entitlement period for unemployment benefits in 30 days, during unemployment zero if claim expired, centered around twelve months
ubendXX	time-varying dummy equal to one if the number of days until expiration of the unemployment claim is smaller or equal to XX days and larger than the number of days referring to the next interval dummy or greater than zero in the last interval (unemployment spells only)
land	10 categories indicating the West German Bundesländer (place of residence): 1 SH, 2 HH, 3 NI, 4 HB, 5 NW, 6 HE, 7 RP, 8 BW, 9 BY, 10 SL
area	West German Bundesländer aggregated into 4 categories (place of residence): 1 SH, NI, HB, HH; 2 NW, 3 HE, RP, SL; 4 BY, BW

&lt;continued on next page&gt;

Table E1: Variable Definitions &lt;continued&gt;

Name	Definition
rtyp	classification of the districts of residence according to their local labor market conditions into 5 groups: 1 regions with very bad labor market conditions (applies to East Germany only), 2 metropolitan areas with high unemployment, 3 regions with medium unemployment, 4 metropolitan areas with very good labor market conditions, 5 non-metropolitan areas with very good labor market conditions
q1, q2, q3, q4	dummy equal to one if spell starts in the first, second, third, fourth quarter of the year
y1999-y2004	year of starting date of spell
cy2000-cy2001, cy2002-2003-2004	time-varying dummies indicating the current calendar year
season1-season4	time-varying dummies indicating the current calendar quarter
ln duruelag, ln duruelagsq	(square of) log duration of last unemployment spell (employment spells only)
longue6, longue12	dummy equal to one if duration of last unemployment spell exceeds six and twelve months, respectively (earnings only)
cycle2, cycle3	dummies indicating second and higher order (third or more) unemployment-employment cycles per person, respectively
dstlag, dstlag2	dummy indicating lagged participation in short-term training during the last (before the last) unemployment-employment cycle
dltlag, dltlag2	dummy indicating lagged participation in short-term training during the last (before the last) unemployment-employment cycle
pst	time-varying dummy equal to one if participation in short-term training has started (unemployment spells only)
stend_XX	time-varying dummy variable equal to one if the number of days until scheduled program end of short-term training is smaller or equal to XX days and larger than the number of days referring to the next interval dummy; the last interval ends at the scheduled completion date (unemployment spells only)
stendXX	time-varying dummy variable equal to one if the number of days after scheduled program completion of short-term training is greater or equal to XX days and smaller than the number of days referring to the next interval dummy; the last interval is open ended (unemployment spells only)
plt	time-varying dummy equal to one if participation in long-term training has started (unemployment spells only)
ltend_XX	time-varying dummy variable equal to one if the number of days until scheduled program end of long-term training is smaller or equal to XX days and larger than the number of days referring to the next interval dummy; the last interval ends at the scheduled completion date (unemployment spells only)

&lt;continued on next page&gt;

Table E1: Variable Definitions &lt;continued&gt;

Name	Definition
ltendXX	time-varying dummy variable equal to one if the number of days after scheduled program completion of long-term training is greater or equal to XX days and smaller than the number of days referring to the next interval dummy; the last interval is open ended (unemployment spells only)
lntst, lntstsq	time-varying, (square of) log of days unemployed at start of short-term training, positive if program has started otherwise zero, centered around 110 days (unemployment spells only)
lntlt, lntltsq	time-varying, (square of) log of days unemployed at start of long-term training, positive if program has started otherwise zero, centered around 110 days (unemployment spells only)
dst	dummy equal to one if participated in short-term training during previous unemployment spell (employment spells only)
dlt	dummy equal to one if participated in long-term training during previous unemployment spell (employment spells only)

Notes: If not noted otherwise, variables are time constant and refer to the start of a spell. In unemployment spells, job characteristics refer to the previous employment. Descriptions of additional interaction terms and aggregated categories are omitted if the content can be inferred from the variable name and the context. For example the variable industry67 is equal to one if industry is either category 6 or 7. Variable names starting with f\_ (fo\_) indicate an interaction with the dummy female (foreign).

## F Detailed Estimation Results

Table F1: Estimated Coefficients

	Model 1 (Employment)	Model 2 (Earnings)
	Unemployment to Employment	
bhaz27	1.868 (0.034)***	2.120 (0.031)***
f_bhaz27	0.388 (0.043)***	
fo_bhaz27	-0.271 (0.046)***	
bhaz32	1.087 (0.033)***	1.051 (0.037)***
bhaz40	1.411 (0.033)***	1.417 (0.040)***
f_bhaz40	0.198 (0.043)***	
fo_bhaz40	-0.213 (0.048)***	
bhaz48	1.133 (0.039)***	0.936 (0.039)***
f_bhaz48	-0.119 (0.052)**	
fo_bhaz48	-0.086 (0.047)*	
bhaz58	1.957 (0.033)***	1.921 (0.032)***
f_bhaz58	0.213 (0.038)***	
fo_bhaz58	-0.294 (0.042)***	
bhaz65	1.592 (0.030)***	1.138 (0.030)***
f_bhaz65	-0.325 (0.038)***	

<continued on next page>

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
fo_bhaz65	-0.317 (0.037)***	
bhaz85	2.214 (0.032)***	1.852 (0.033)***
f_bhaz85	-0.063 (0.038)*	
fo_bhaz85	-0.486 (0.040)***	
bhaz94	1.794 (0.032)***	1.190 (0.035)***
f_bhaz94	-0.452 (0.047)***	
fo_bhaz94	-0.399 (0.041)***	
bhaz110	1.927 (0.035)***	1.436 (0.034)***
f_bhaz110	-0.089 (0.039)**	
fo_bhaz110	-0.392 (0.041)***	
bhaz124	1.543 (0.036)***	0.980 (0.039)***
f_bhaz124	-0.314 (0.047)***	
fo_bhaz124	-0.288 (0.042)***	
bhaz144	1.554 (0.033)***	1.143 (0.038)***
fo_bhaz144	-0.234 (0.040)***	
bhaz166	1.488 (0.037)***	1.139 (0.039)***
f_bhaz166	0.091 (0.045)**	
fo_bhaz166	-0.217 (0.043)***	
bhaz184	1.129 (0.041)***	0.766 (0.039)***
f_bhaz184	-0.058 (0.043)	
fo_bhaz184	-0.089 (0.048)*	
bhaz214	0.893 (0.027)***	0.591 (0.035)***
bhaz306	0.713 (0.040)***	0.342 (0.049)***
fo_bhaz306	-0.130 (0.051)**	
bhaz366	0.418 (0.045)***	0.175 (0.043)***
f_bhaz366	0.161 (0.049)***	
fo_bhaz366	-0.071 (0.051)	
bhaz470	0.240 (0.050)***	-0.138 (0.051)***
f_bhaz470	0.096 (0.054)*	
fo_bhaz470	-0.120 (0.053)**	
bhaz622	-0.176 (0.044)***	-0.668 (0.049)***
f_bhaz622	0.212 (0.064)***	
bhaz955	-0.594 (0.055)***	
female	0.088 (0.052)*	-0.023 (0.061)
foreign	0.058 (0.021)***	-0.139 (0.018)***
ethnicgerman	0.117 (0.032)***	0.101 (0.039)***
agegroup2	-0.080 (0.016)***	-0.087 (0.018)***
agegroup3	-0.156 (0.017)***	-0.151 (0.020)***
agegroup4	-0.249 (0.020)***	-0.268 (0.024)***
agegroup5	-0.318 (0.021)***	-0.342 (0.024)***
agegroup6	-0.455 (0.027)***	-0.550 (0.037)***
f_agegroup4	0.075 (0.029)***	0.084 (0.034)**
f_agegroup6	-0.138 (0.039)***	-0.154 (0.046)***

&lt;continued on next page&gt;

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
schooling3	0.101 (0.019)***	0.087 (0.023)***
schooling4	0.100 (0.026)***	0.050 (0.032)
education3	0.154 (0.017)***	0.151 (0.021)***
education4	0.155 (0.032)***	0.186 (0.038)***
f_education34	-0.055 (0.024)**	-0.072 (0.029)**
kids	0.076 (0.017)***	0.080 (0.021)***
f_youngchild	-0.341 (0.029)***	-0.353 (0.032)***
f_kids	-0.088 (0.027)***	-0.142 (0.035)***
family3	-0.065 (0.026)**	-0.060 (0.035)*
family4	-0.036 (0.028)	-0.019 (0.037)
family5	0.012 (0.014)	-0.027 (0.017)
health3	-0.375 (0.023)***	-0.401 (0.030)***
health4	-0.567 (0.025)***	-0.606 (0.032)***
disabled	-0.031 (0.007)***	-0.030 (0.008)***
taredu3	0.070 (0.014)***	0.078 (0.016)***
tarexp	-0.091 (0.017)***	-0.150 (0.022)***
prevtrans	0.159 (0.012)***	0.122 (0.014)***
endlastjob3	-0.101 (0.024)***	-0.084 (0.024)***
endlastjob4	0.071 (0.014)***	0.040 (0.019)**
seasonwork	0.302 (0.024)***	0.240 (0.027)***
whitecollar	-0.140 (0.029)***	-0.101 (0.030)***
bluecollar	-0.033 (0.028)	-0.017 (0.031)
parttime	-0.124 (0.028)***	-0.135 (0.028)***
lwquart2	0.071 (0.017)***	0.061 (0.021)***
lwquart3	0.148 (0.018)***	0.124 (0.022)***
lwquart4	0.270 (0.020)***	0.226 (0.024)***
occupation2	0.141 (0.077)*	0.054 (0.101)
occupation3	0.547 (0.084)***	0.347 (0.107)***
occupation4	0.342 (0.078)***	0.254 (0.104)**
occupation5	0.310 (0.078)***	0.210 (0.102)**
occupation6	0.209 (0.078)***	0.150 (0.102)
occupation7	0.107 (0.081)	0.046 (0.101)
occupation8	0.150 (0.080)*	0.101 (0.103)
occupation9	0.232 (0.082)***	0.163 (0.103)
occupation10	0.072 (0.090)	0.042 (0.110)
f_occupation4	-0.298 (0.041)***	-0.330 (0.050)***
f_occupation5	-0.209 (0.042)***	-0.221 (0.055)***
f_occupation8	-0.024 (0.037)	-0.020 (0.041)
fo_occupation3	-0.323 (0.058)***	-0.301 (0.079)***
industry2	-0.598 (0.183)***	-0.003 (0.301)
industry3	-0.734 (0.178)***	-0.204 (0.291)
industry4	-0.412 (0.178)**	0.083 (0.294)
industry5	-0.589 (0.178)***	-0.108 (0.293)

&lt;continued on next page&gt;



Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
industry67	-0.671 (0.177)***	-0.213 (0.292)
f_industry3	0.123 (0.052)**	0.149 (0.064)**
f_industry56	0.112 (0.047)**	0.221 (0.056)***
f_industry7	0.194 (0.050)***	0.278 (0.057)***
fo_industry4	-0.109 (0.025)***	-0.114 (0.032)***
fo_industry6	0.136 (0.026)***	0.159 (0.032)***
hasclaim	-0.046 (0.015)***	-0.073 (0.019)***
ubend60	0.174 (0.025)***	0.206 (0.038)***
ubend30	0.349 (0.027)***	0.368 (0.039)***
totclaim	0.065 (0.006)***	0.080 (0.008)***
totclaimsq	-0.002 (0.000)***	-0.003 (0.000)***
sick	-2.265 (0.126)***	-2.270 (0.171)***
eversick	-0.149 (0.024)***	-0.117 (0.029)***
land5	-0.127 (0.014)***	-0.090 (0.016)***
land8	-0.007 (0.020)	0.027 (0.022)
land9	0.246 (0.014)***	0.244 (0.017)***
land10	-0.178 (0.044)***	-0.195 (0.053)***
rtyp2	-0.131 (0.018)***	-0.151 (0.019)***
rtyp5	0.144 (0.014)***	0.139 (0.016)***
season2	0.372 (0.012)***	0.252 (0.014)***
season3	0.194 (0.013)***	0.124 (0.017)***
season4	-0.272 (0.014)***	-0.255 (0.017)***
y1999q4	0.047 (0.025)*	0.087 (0.027)***
y2000q1	0.044 (0.026)*	0.090 (0.027)***
y2000q2	-0.127 (0.030)***	-0.115 (0.029)***
y2000q3	-0.095 (0.030)***	-0.091 (0.029)***
y2000q4	-0.148 (0.027)***	-0.123 (0.029)***
y2001q1	-0.174 (0.025)***	-0.135 (0.028)***
y2001q2	-0.358 (0.029)***	-0.297 (0.032)***
y2001q3	-0.337 (0.027)***	-0.263 (0.031)***
y2001q4	-0.377 (0.025)***	-0.248 (0.030)***
y2002q1	-0.415 (0.034)***	
y2002q2	-0.794 (0.043)***	
y2002q3	-0.710 (0.039)***	
y2002q4	-0.601 (0.032)***	
y2003q1	-0.583 (0.037)***	
y2003q2	-0.969 (0.045)***	
y2003q3	-0.769 (0.043)***	
y2003q4	-0.685 (0.037)***	
y2004q1	-0.702 (0.038)***	
y2004q2	-1.039 (0.058)***	
y2004q3	-1.073 (0.058)***	
y2004q4	-1.604 (0.096)***	

&lt;continued on next page&gt;

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
cycle2	0.192 (0.016)***	
cycle3	0.458 (0.021)***	
dstlag	-0.175 (0.034)***	
dstlag2	-0.055 (0.063)	
dltlag	-0.120 (0.047)**	
dltlag2	-0.171 (0.072)**	
pst	0.057 (0.123)	0.324 (0.242)
stend_16	-1.175 (0.074)***	-1.125 (0.098)***
stend_15	-0.202 (0.058)***	-0.214 (0.080)***
stend31	-0.592 (0.044)***	-0.610 (0.067)***
stend361	-0.789 (0.069)***	-0.904 (0.101)***
f_stend31	0.186 (0.072)**	0.287 (0.091)***
f_stend361	0.179 (0.110)	0.301 (0.137)**
lntst	-0.006 (0.055)	0.135 (0.079)*
lntstsq	0.009 (0.007)	-0.011 (0.010)
pst_female	-0.113 (0.057)**	-0.048 (0.068)
pst_foreign	0.020 (0.038)	0.111 (0.051)**
pst_disabled	0.089 (0.041)**	0.068 (0.039)*
pst_health3	0.075 (0.064)	-0.103 (0.093)
pst_health4	0.332 (0.078)***	0.353 (0.097)***
pst_age	0.370 (0.212)*	0.253 (0.282)
pst_agesq	-0.055 (0.026)**	-0.038 (0.036)
pst_occupation2	0.400 (0.122)***	0.480 (0.182)***
pst_occupation45	0.442 (0.121)***	0.527 (0.171)***
pst_occupation6	0.367 (0.125)***	0.458 (0.181)**
pst_occupation9	0.261 (0.141)*	0.203 (0.190)
pst_occupation7810	0.460 (0.118)***	0.476 (0.175)***
plt	0.544 (0.072)***	1.751 (0.177)***
ltend_361	-3.227 (0.124)***	-4.373 (0.194)***
ltend_360	-2.654 (0.100)***	-3.743 (0.168)***
ltend_180	-2.173 (0.086)***	-3.110 (0.145)***
ltend_90	-1.511 (0.081)***	-2.164 (0.124)***
ltend_30	-0.513 (0.059)***	-0.934 (0.100)***
ltend91	-0.174 (0.058)***	0.078 (0.083)
ltend181	-0.366 (0.064)***	0.012 (0.101)
ltend361	-0.468 (0.090)***	-0.107 (0.139)
ltend541	-0.326 (0.097)***	-0.119 (0.129)
fo_ltend_180	0.301 (0.145)**	0.515 (0.163)***
fo_ltend_90	0.305 (0.136)**	0.424 (0.148)***
lntlt	-0.059 (0.052)	0.165 (0.081)**
lntltsq	0.015 (0.007)**	-0.029 (0.012)**
plt_female	0.056 (0.049)	0.190 (0.065)***
plt_foreign	-0.078 (0.065)	0.062 (0.090)

&lt;continued on next page&gt;

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
plt_disabled	0.034 (0.021)	0.034 (0.027)
plt_health3	0.123 (0.083)	0.151 (0.118)
plt_health4	0.255 (0.082)***	0.219 (0.135)
plt_sick	-0.901 (1.003)	-0.521 (1.027)
plt_eversick	0.090 (0.073)	0.021 (0.101)
plt_age	0.988 (0.255)***	1.049 (0.412)**
plt_agesq	-0.135 (0.033)***	-0.147 (0.053)***
plt_education3	-0.236 (0.050)***	-0.272 (0.079)***
plt_foeducation3	0.106 (0.089)	0.274 (0.122)**
plt_foeducation4	-0.328 (0.195)*	-0.361 (0.320)
plt_occupation45	-0.052 (0.050)	-0.006 (0.082)
plt_occupation78910	0.099 (0.049)**	0.017 (0.074)
Intercept	-6.269 (0.196)***	-6.841 (0.304)***
Unemployment to Short-Term Training		
bhaz17	-0.227 (0.055)***	-0.149 (0.072)**
fo_bhaz17	0.229 (0.108)**	
bhaz37	-0.515 (0.068)***	-0.291 (0.068)***
f_bhaz37	0.239 (0.082)***	
fo_bhaz37	0.402 (0.110)***	
bhaz66	-0.428 (0.070)***	-0.216 (0.073)***
f_bhaz66	0.195 (0.079)**	
fo_bhaz66	0.327 (0.105)***	
bhaz99	-0.335 (0.063)***	-0.177 (0.074)**
fo_bhaz99	0.335 (0.111)***	
bhaz140	-0.425 (0.056)***	-0.155 (0.064)**
fo_bhaz140	0.349 (0.088)***	
c_bhaz140	0.215 (0.052)***	
bhaz568	-0.701 (0.092)***	-0.329 (0.087)***
f_bhaz568	-0.184 (0.083)**	
fo_bhaz568	0.393 (0.112)***	
c_bhaz568	0.481 (0.079)***	
female	-0.066 (0.032)**	-0.027 (0.039)
foreign	-0.520 (0.088)***	-0.286 (0.054)***
agegroup2	-0.012 (0.039)	-0.017 (0.052)
agegroup3	-0.038 (0.040)	-0.008 (0.054)
agegroup4	-0.114 (0.040)***	-0.007 (0.058)
agegroup5	-0.218 (0.041)***	-0.140 (0.059)**
agegroup6	-0.663 (0.053)***	-0.566 (0.076)***
fo_schooling3	0.089 (0.043)**	0.133 (0.055)**
education23	-0.338 (0.177)*	-0.732 (0.204)***
education4	-0.392 (0.188)**	-0.664 (0.212)***
family4	0.195 (0.049)***	0.210 (0.064)***
family5	-0.153 (0.025)***	-0.182 (0.035)***

&lt;continued on next page&gt;

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
health3	-0.075 (0.044)*	-0.126 (0.060)**
health4	-0.174 (0.045)***	-0.199 (0.056)***
disabled	0.125 (0.027)***	0.137 (0.031)***
seekpt	-0.196 (0.045)***	-0.209 (0.052)***
taredu3	0.078 (0.027)***	0.006 (0.035)
taredu4	0.148 (0.078)*	0.131 (0.095)
tarexp	-0.029 (0.038)	0.050 (0.058)
earlycontact	0.142 (0.026)***	0.125 (0.037)***
seasonwork	-0.141 (0.041)***	-0.088 (0.061)
lwquart2	0.068 (0.036)*	0.043 (0.050)
lwquart3	0.138 (0.035)***	0.118 (0.051)**
lwquart4	-0.044 (0.038)	-0.009 (0.053)
occupation3	-0.093 (0.085)	-0.022 (0.126)
occupation4	-0.025 (0.044)	0.008 (0.055)
occupation5	-0.006 (0.041)	0.062 (0.056)
occupation6	0.102 (0.043)**	0.128 (0.064)**
occupation7	0.141 (0.043)***	0.196 (0.055)***
occupation8	0.094 (0.049)*	0.103 (0.067)
occupation9	-0.078 (0.095)	-0.186 (0.129)
occupation10	-0.330 (0.137)**	-0.203 (0.165)
f_occupation9	0.133 (0.103)	0.032 (0.137)
f_occupation10	0.449 (0.239)*	0.443 (0.273)
industry3	0.279 (0.043)***	0.313 (0.062)***
industry5	0.216 (0.046)***	0.261 (0.063)***
industry6	0.213 (0.046)***	0.339 (0.070)***
industry7	0.085 (0.049)*	0.221 (0.070)***
f_industry4	0.222 (0.139)	0.362 (0.153)**
hasclaim	0.188 (0.028)***	0.195 (0.040)***
totclaim	0.027 (0.013)**	0.058 (0.019)***
totclaimsq	-0 (0.000)	-0.001 (0.001)*
sick	-0.191 (0.097)**	-0.139 (0.152)
eversick	0.107 (0.044)**	0.028 (0.060)
offben	-0.145 (0.145)	-0.081 (0.192)
eversanc	0.370 (0.066)***	0.409 (0.104)***
rtyp3	-0.061 (0.031)**	-0.110 (0.041)***
rtyp4	-0.177 (0.045)***	-0.266 (0.054)***
rtyp5	-0.108 (0.035)***	-0.100 (0.047)**
cy2000	0.838 (0.099)***	
cy2001	0.949 (0.106)***	
cy200220032004	1.169 (0.117)***	
season2	0.286 (0.032)***	0.145 (0.042)***
season3	0.191 (0.036)***	-0.056 (0.043)
season4	-0.010 (0.033)	-0.214 (0.045)***

&lt;continued on next page&gt;

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
q2	0.067 (0.030)**	0.008 (0.046)
q3	0.153 (0.027)***	0.110 (0.037)***
y20002001	0.029 (0.055)	0.403 (0.045)***
y2002	0.195 (0.077)**	
y20032004	0.360 (0.077)***	
cycle2	-0.207 (0.044)***	
cycle3	-0.387 (0.057)***	
Intercept	-8.032 (0.224)***	-7.295 (0.265)***
Unemployment to Long-Term Training		
bhaz78	0.055 (0.055)	-0.035 (0.061)
bhaz114	0.253 (0.038)***	0.191 (0.053)***
bhaz336	0.361 (0.065)***	0.380 (0.092)***
bhaz428	-0.513 (0.062)***	-0.413 (0.099)***
female	-0.307 (0.098)***	-0.348 (0.136)**
foreign	-0.244 (0.038)***	-0.281 (0.049)***
ethnicgerman	0.375 (0.081)***	0.453 (0.111)***
agegroup2	-0.073 (0.044)*	-0.105 (0.057)*
agegroup34	-0.179 (0.043)***	-0.182 (0.054)***
agegroup5	-0.445 (0.054)***	-0.451 (0.073)***
agegroup6	-0.991 (0.064)***	-1.024 (0.081)***
schooling2	-0.822 (0.173)***	-1.186 (0.234)***
schooling3	-0.607 (0.172)***	-0.973 (0.229)***
schooling4	-0.162 (0.183)	-0.508 (0.230)**
f_schooling3	0.332 (0.100)***	0.378 (0.138)***
f_schooling4	0.087 (0.118)	0.072 (0.162)
education2	-0.231 (0.178)	-0.451 (0.230)*
education3	-0.228 (0.176)	-0.473 (0.225)**
education4	-0.358 (0.194)*	-0.597 (0.253)**
f_youngchild	-0.157 (0.058)***	-0.163 (0.078)**
family3	0.171 (0.065)***	0.306 (0.099)***
family4	0.469 (0.060)***	0.582 (0.084)***
family5	0.068 (0.032)**	0.129 (0.042)***
health4	-0.121 (0.052)**	-0.105 (0.082)
seekpt	-0.356 (0.054)***	-0.404 (0.065)***
taredu4	-0.125 (0.077)	-0.219 (0.094)**
tarexp	-0.302 (0.043)***	-0.350 (0.064)***
earlycontact	0.329 (0.032)***	0.412 (0.039)***
prevtrans	-0.217 (0.029)***	-0.316 (0.040)***
lwquart2	0.139 (0.049)***	0.105 (0.070)
lwquart3	0.277 (0.064)***	0.259 (0.096)***
lwquart4	0.274 (0.093)***	0.223 (0.131)*
lnlwage	0.538 (0.118)***	0.639 (0.170)***
lnlwagesq	-0.098 (0.023)***	-0.112 (0.034)***

&lt;continued on next page&gt;

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
seasonwork	-0.217 (0.057)***	-0.226 (0.081)***
bluecollar	-0.102 (0.043)**	-0.106 (0.059)*
whitecollar	0.163 (0.039)***	0.111 (0.055)**
occupation2	0.173 (0.089)*	0.153 (0.115)
occupation4	0.235 (0.093)**	0.230 (0.118)*
occupation5	0.127 (0.090)	0.079 (0.115)
occupation6	0.126 (0.088)	0.155 (0.114)
occupation7	0.488 (0.085)***	0.588 (0.111)***
occupation8	0.259 (0.088)***	0.313 (0.112)***
occupation9	0.169 (0.096)*	0.253 (0.123)**
industry3	0.117 (0.033)***	0.138 (0.052)***
industry4	-0.492 (0.061)***	-0.516 (0.088)***
industry7	-0.262 (0.049)***	-0.295 (0.063)***
f_industry4	0.514 (0.140)***	0.513 (0.179)***
fo_industry7	0.260 (0.080)***	0.291 (0.104)***
hasclaim	0.182 (0.044)***	0.194 (0.059)***
ubend120	0.103 (0.063)	0.055 (0.072)
ubend90	0.238 (0.063)***	0.218 (0.074)***
ubend60	0.084 (0.050)*	0.089 (0.065)
offben	-0.615 (0.205)***	-0.976 (0.336)***
land2	0.275 (0.079)***	0.346 (0.103)***
land3	0.070 (0.045)	0.048 (0.059)
land6	0.043 (0.058)	0.122 (0.077)
land7	0.136 (0.063)**	0.133 (0.086)
land8	0.091 (0.058)	0.222 (0.082)***
land9	0.065 (0.046)	0.111 (0.068)
rtyp2	0.061 (0.047)	0.090 (0.061)
rtyp5	0.123 (0.041)***	0.128 (0.058)**
season23	-0.048 (0.030)	-0.099 (0.037)***
season4	-0.564 (0.040)***	-0.589 (0.049)***
q3	-0.029 (0.035)	0.036 (0.050)
q4	-0.157 (0.035)***	-0.091 (0.046)**
y2000	-0.074 (0.044)*	-0.015 (0.054)
y2001	-0.394 (0.046)***	-0.385 (0.056)***
y2002	-0.964 (0.066)***	
y2003	-1.367 (0.078)***	
y2004	-1.625 (0.113)***	
pst	0.289 (0.042)***	0.236 (0.056)***
Intercept	-7.285 (0.228)***	-6.462 (0.355)***
Employment / Earnings		
bhaz14		0.498 (0.027)***
bhaz24		1.053 (0.023)***
bhaz41		1.569 (0.027)***

&lt;continued on next page&gt;

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
bhaz47	0.424 (0.025)***	
f_bhaz47	-0.205 (0.038)***	
bhaz49		1.941 (0.029)***
bhaz56		2.529 (0.028)***
bhaz72		3.197 (0.031)***
bhaz89	0.431 (0.030)***	
fo_bhaz89	0.060 (0.036)*	
bhaz133	0.321 (0.028)***	
bhaz162	0.584 (0.031)***	
f_bhaz162	0.186 (0.045)***	
bhaz184	0.201 (0.037)***	
fo_bhaz184	0.096 (0.050)*	
bhaz212	0.625 (0.032)***	
fo_bhaz212	-0.204 (0.050)***	
bhaz236	1.117 (0.037)***	
f_bhaz236	-0.085 (0.050)*	
fo_bhaz236	-0.454 (0.052)***	
bhaz254	1.332 (0.039)***	
f_bhaz254	-0.569 (0.052)***	
fo_bhaz254	-0.592 (0.053)***	
bhaz274	1.805 (0.036)***	
f_bhaz274	-0.812 (0.059)***	
fo_bhaz274	-0.934 (0.058)***	
bhaz290	1.699 (0.037)***	
f_bhaz290	-0.727 (0.053)***	
fo_bhaz290	-0.996 (0.056)***	
bhaz311	1.007 (0.036)***	
f_bhaz311	-0.228 (0.042)***	
fo_bhaz311	-0.554 (0.050)***	
bhaz366	0.044 (0.049)	
f_bhaz366	-0.105 (0.051)**	
fo_bhaz366	-0.112 (0.058)*	
bhaz441	-0.022 (0.043)	
f_bhaz441	-0.110 (0.047)**	
fo_bhaz441	-0.107 (0.050)**	
bhaz580	0.228 (0.039)***	
f_bhaz580	-0.193 (0.045)***	
fo_bhaz580	-0.309 (0.051)***	
bhaz732	-0.326 (0.038)***	
fo_bhaz732	-0.236 (0.051)***	
bhaz992	-0.371 (0.035)***	
fo_bhaz992	-0.253 (0.050)***	
female	-0.171 (0.040)***	0.404 (0.023)***

&lt;continued on next page&gt;

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
foreign	0.293 (0.035)***	0.113 (0.016)***
ethnicgerman	-0.099 (0.034)***	0.092 (0.040)**
age	-0.452 (0.086)***	-0.447 (0.109)***
agesq	0.071 (0.011)***	0.055 (0.014)***
f_age	-0.393 (0.150)***	0.361 (0.171)**
f_agesq	0.037 (0.019)**	-0.037 (0.022)*
education2	0.171 (0.018)***	0.356 (0.097)***
education3		0.151 (0.095)
education4	-0.239 (0.055)***	-0.245 (0.099)**
f_education3	-0.048 (0.024)**	
f_education4	0.221 (0.077)***	
schooling2	0.343 (0.064)***	0.180 (0.093)*
schooling3	0.278 (0.063)***	0.109 (0.091)
schooling4	0.234 (0.074)***	-0.138 (0.093)
f_schooling4	0.059 (0.042)	
fo_schooling3	0.050 (0.034)	
fo_schooling4	0.078 (0.045)*	
kids	-0.058 (0.015)***	-0.027 (0.017)
family3	0.065 (0.031)**	0.079 (0.036)**
family4	0.040 (0.032)	0.114 (0.038)***
family5	-0.055 (0.015)***	-0.007 (0.017)
health	0.120 (0.018)***	0.152 (0.023)***
eversick	0.078 (0.028)***	0.214 (0.029)***
eversanc		0.168 (0.055)***
lnlwage	0.022 (0.034)	
lnlwagesq	-0.009 (0.007)	
lwquart2		-0.357 (0.024)***
lwquart3	-0.074 (0.017)***	-0.777 (0.023)***
lwquart4	-0.113 (0.023)***	-1.419 (0.028)***
endlastjob2	0.046 (0.016)***	
endlastjob3	-0.082 (0.031)***	
endlastjob4	0.157 (0.020)***	
totclaim	-0.014 (0.001)***	
bluecollar	0.151 (0.018)***	0.261 (0.021)***
occupation2	0.051 (0.040)	0.574 (0.043)***
occupation3		0.557 (0.053)***
occupation4	0.025 (0.043)	0.399 (0.041)***
occupation5	0.063 (0.041)	0.400 (0.039)***
occupation6	0.017 (0.046)	0.544 (0.039)***
occupation7	-0.016 (0.045)	0.335 (0.038)***
occupation8	-0.097 (0.049)**	
occupation9	-0.245 (0.066)***	
occupation10	-0.333 (0.083)***	

&lt;continued on next page&gt;



Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
occupation910		-0.273 (0.047)***
f_occupation3	0.330 (0.064)***	
f_occupation4	0.101 (0.048)**	
f_occupation8	0.041 (0.045)	0.181 (0.042)***
f_occupation9	0.195 (0.069)***	
f_occupation10	0.408 (0.133)***	
f_occupation910		0.295 (0.047)***
fo_occupation8		-0.099 (0.051)*
fo_occupation910		-0.101 (0.060)*
seasonwork	0.339 (0.040)***	
industry2	-0.129 (0.052)**	
industry3	-0.140 (0.043)***	
industry4	0.808 (0.038)***	
industry5	-0.167 (0.036)***	
industry6	0.008 (0.037)	
f_industry3	0.189 (0.043)***	
f_industry5	0.239 (0.034)***	
f_industry6	0.148 (0.040)***	
land1	0.123 (0.025)***	0.190 (0.032)***
land2	0.119 (0.036)***	0.064 (0.042)
land3	0.073 (0.018)***	0.153 (0.020)***
land4	0.060 (0.052)	0.247 (0.062)***
land6	-0.034 (0.023)	0.021 (0.024)
land8	-0.102 (0.020)***	-0.049 (0.024)**
land9	0.063 (0.016)***	0.011 (0.019)
season23	-0.452 (0.075)***	
season4	0.497 (0.030)***	
s2_seasonwork	-0.621 (0.059)***	
s3_seasonwork	-0.519 (0.053)***	
s4_seasonwork	0.179 (0.044)***	
s2_industry3	0.196 (0.080)**	
s2_industry4	-1.163 (0.088)***	
s2_industry567	0.287 (0.076)***	
s3_industry3	0.278 (0.082)***	
s3_industry4	-1.036 (0.086)***	
s3_industry5	0.347 (0.080)***	
s3_industry6	0.399 (0.078)***	
s3_industry7	0.573 (0.077)***	
s4_industry3	-0.062 (0.046)	
s4_industry4	-0.805 (0.039)***	
s4_industry5	-0.050 (0.034)	
s4_industry6	-0.224 (0.042)***	
q1	-0.131 (0.013)***	

&lt;continued on next page&gt;

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
q3	-0.068 (0.015)***	
y2000	-0.013 (0.033)	
y2001	0.020 (0.033)	
y2002	0.045 (0.036)	
y2003	0.048 (0.038)	
y2004q1	-0.032 (0.052)	
y2004q2	0.178 (0.047)***	
y2004q3	0.357 (0.064)***	
y2004q4	0.325 (0.085)***	
cycle2	0.259 (0.015)***	
cycle3	0.513 (0.021)***	
lnuruelag	0.357 (0.062)***	
lnuruelagsq	-0.034 (0.007)***	
longue6		0.235 (0.020)***
longue12		0.207 (0.029)***
dstlag	-0.187 (0.042)***	
dltlag	-0.211 (0.049)***	
dst	0.258 (0.063)***	0.131 (0.048)***
dst_female	0.045 (0.050)	-0.200 (0.052)***
dst_foreign	-0.084 (0.048)*	-0.153 (0.058)***
dst_ethnicgerman	-0.305 (0.148)**	-0.209 (0.151)
dst_disabled	-0.084 (0.043)*	
dst_health	-0.095 (0.068)	
dst_eversick	0.084 (0.067)	
dst_schooling34	-0.170 (0.063)***	
dst_occupation67	-0.054 (0.055)	
dst_occupation789		0.054 (0.057)
dlt	0.163 (0.071)**	-0.107 (0.127)
dlt_female	-0.170 (0.084)**	-0.185 (0.054)***
dlt_foreign	0.296 (0.102)***	-0.067 (0.057)
dlt_ethnicgerman	0.197 (0.136)	-0.143 (0.143)
dlt_health	0.119 (0.075)	
dlt_eversick	0.224 (0.091)**	
dlt_schooling4	-0.097 (0.080)	
dlt_foschooling34	-0.137 (0.114)	
dlt_education3	-0.157 (0.071)**	
dlt_feducation3	0.182 (0.111)	
dlt_occupation910	0.226 (0.108)**	
dlt_occupation5		0.098 (0.075)
dlt_occupation789		0.283 (0.060)***
dlt_longue6		-0.181 (0.096)*
dlt_longue12		-0.176 (0.064)***
Intercept	-7.658 (0.171)***	-6.130 (0.081)***

&lt;continued on next page&gt;

Table F1: Estimated Coefficients &lt;continued&gt;

	Model 1 (Employment)	Model 2 (Earnings)
Factor Loadings on First Latent Factor ( $w_1$ )		
Unemployment to Employment	-0.779 (0.021)***	0.044 (0.072)
Unemployed, treated with short-term tr.	0.339 (0.085)***	-0.151 (0.215)
Unemployed, treated with long-term tr.	0.307 (0.107)***	0.136 (0.403)
Unem. to short-term training	-0.185 (0.079)**	-0.057 (0.147)
Unem. to long-term training	-0.112 (0.111)	-0.502 (0.350)
Employment to unemployment / earnings	0.127 (0.020)***	0.669 (0.051)***
Employed, treated with short-term tr.	-0.106 (0.077)	
Employed, treated with long-term tr.	-0.253 (0.141)*	
Factor Loadings on Second Latent Factor ( $w_2$ )		
Unemployment to employment	-0.426 (0.020)***	0.329 (0.051)***
Unemployed, treated with short-term tr.	0.110 (0.047)**	0.076 (0.056)
Unemployed, treated with long-term tr.	-0.109 (0.065)*	0.917 (0.100)***
Unem. to short-term training	0	0
Unem. to long-term training	0.036 (0.052)	-1.114 (0.080)***
Employment to unemployment / earnings	0.773 (0.024)***	-0.073 (0.028)***
Employed, treated with short-term tr.	0.159 (0.044)***	
Employed, treated with long-term tr.	0.224 (0.069)***	
Probabilities		
$\Pr(w_1 = 1)$	0.904 (0.009)***	0.958 (0.009)***
$\Pr(w_2 = 1)$	0.151 (0.010)***	0.787 (0.017)***
Model Statistics		
Log-Likelihood	-232,678.02	-114,220.94
# Parameters	524	361
# Spells	318,508	166,567
# Individuals	45,459	45,459

Notes: Model 1 refers to the model with unemployment duration and employment duration as outcomes, model 2 to the model with unemployment duration and daily earnings as outcomes. Bootstrapped standard errors based on 250 replications are in parentheses. \*, \*\* and \*\*\* denote significance at the 10%-, 5%- and 1%- level, respectively, assuming asymptotic normality of the corresponding coefficient.

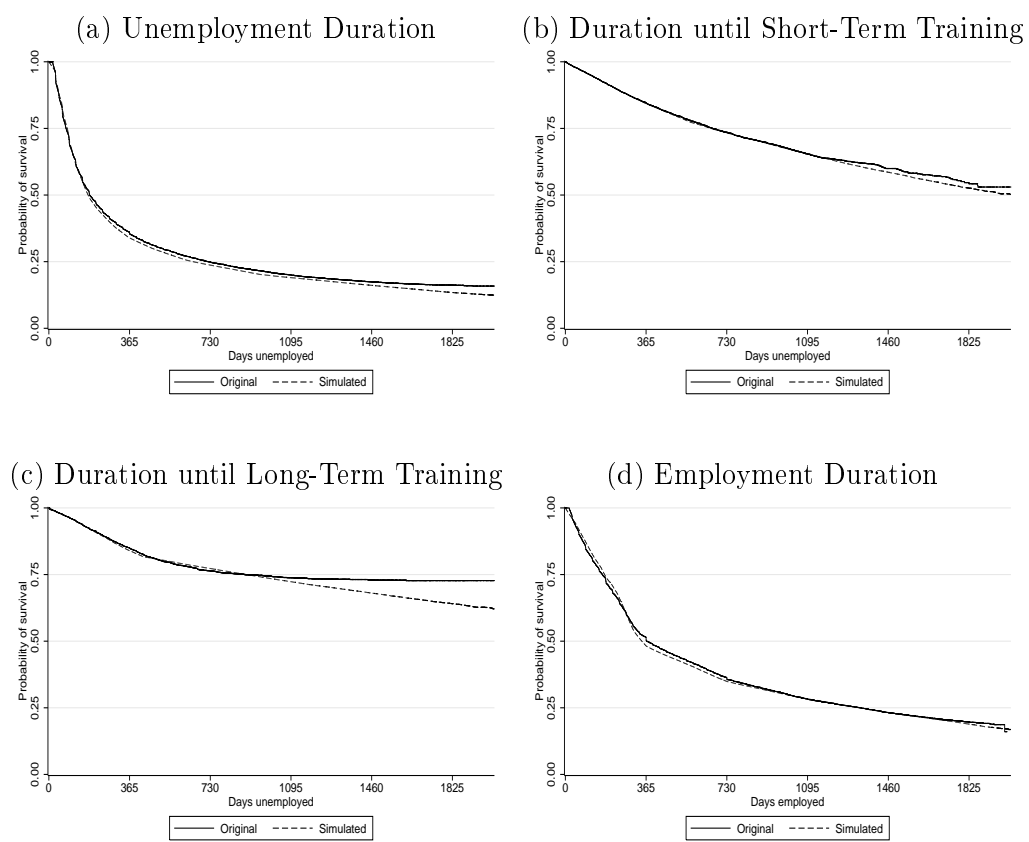
## G Model Fit

In order to assess the fit of the estimated models, we simulate ten unemployment-employment cycles for each person in the inflow sample, each time assigning a new draw from the unobserved heterogeneity distribution and a randomly drawn percentile of the duration distribution. The percentiles are drawn from a uniform distribution on  $[0, 1)$ . The unobserved heterogeneity terms are drawn from their estimated distribution. In order to find for each observation in the sample and for a given percentile  $q$  the corresponding survival time  $t(q) = \inf\{t : S(t) \leq q\}$ , with  $S(t)$  the survivor function, we proceed as follows. First, we predict the survival probability at the beginning,  $S(\underline{t}_d)$ , and the end,  $S(\bar{t}_d)$ , of each segment  $d$  of our piecewise constant exponential model using the estimated coefficients on the observed variables and the randomly drawn value from the unobserved heterogeneity distribution. Second, we determine the segment in which  $S(\underline{t}_d) < 1 - q \leq S(\bar{t}_d)$ . Third, we compute the survival time  $t(q) = \underline{t}_d - \ln[(1 - q)/S(\bar{t}_d)]/\theta_d$ , where  $\theta_d$  is the exponential hazard rate in the  $d$ th segment.

We first simulate the waiting times until short-term training and long-term training, respectively. Then we assign each person a randomly drawn planned program duration. We draw the planned program durations from the distribution of planned program durations in our sample. Next, we calculate the survival time in unemployment conditional on the simulated starting date and length of training. Finally, we compute the subsequent employment duration and daily earnings in an analogous way, using the observations following simulated unemployment spells that end with a transition to employment. Simulated waiting times until treatment as well as unemployment and employment durations are censored at the end of the observation period in the original data, i.e. December 31, 2004, and daily earnings at the social security threshold in 2000, i.e. €144.

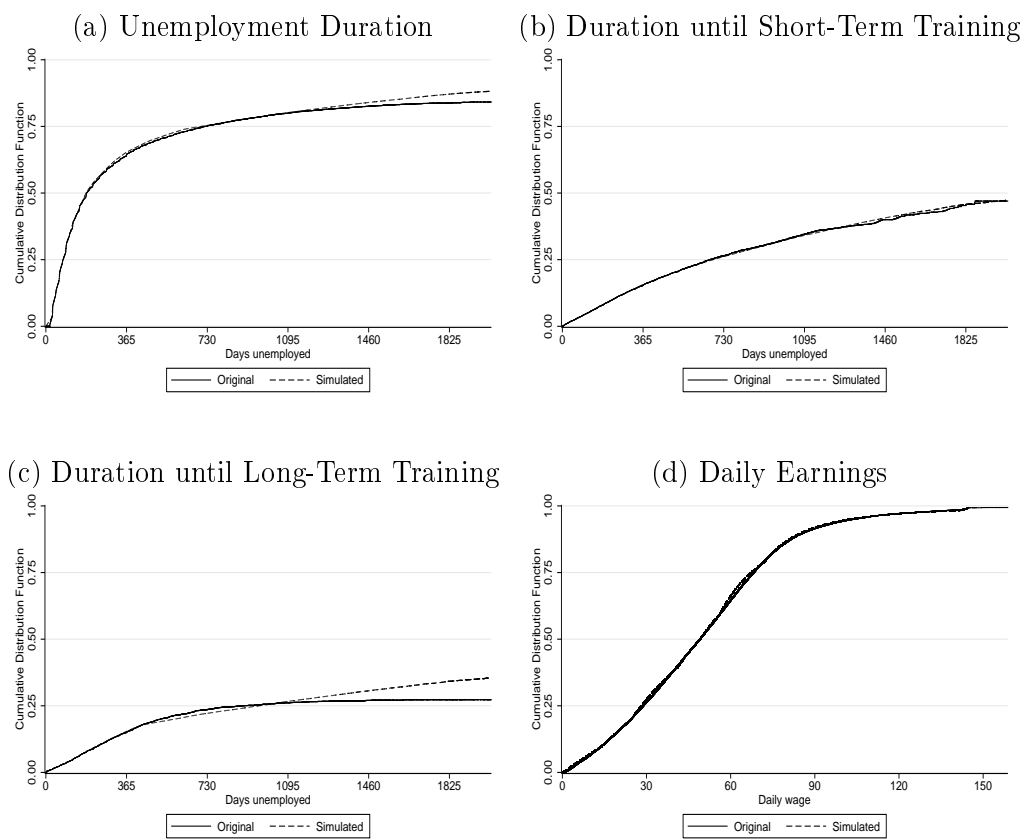
Figure G1 displays the simulated and original distributions for the model with employment duration as the second outcome and figure G2 the simulated and original distributions for the model with daily earnings as the second outcome. The model fit is in general excellent. In regions in which there are many exits, actual and simulated survivor functions nearly coincide. For the durations until employment and until long-term training (panel (a) and (c)), respectively, the simulated survivor functions slightly overestimate the exit probability at very long durations where the actual probability of exit is extremely small and thus the survivor almost horizontal.

Figure G1: Fit of the Model with Employment Duration as the Second Outcome



Notes: Kaplan-Meier (1958) estimates of the survival probabilities are based on ten simulated unemployment-employment cycles for each person in the sample.

Figure G2: Fit of the Model with Daily Earnings as the Second Outcome



Notes: Kaplan-Meier (1958) estimates of the cumulative distribution functions are based on ten simulated unemployment-employment cycles for each person in the sample.

## H Simulation of Treatment and Nontreatment Outcomes

We obtain treatment effects by simulating treatment and nontreatment outcomes for a subset of the originally treated persons in the data. Specifically, we consider those persons in our sample who start training within the first 731 days of their first unemployment spell. We exclude later training starts in order to avoid extrapolations beyond the time horizon of our data. We simulate ten unemployment-employment cycles for each person, each time assigning a new draw from the unobserved heterogeneity distribution and a randomly drawn percentile of the duration distribution. The simulation procedure is analogous to that outlined in section G except that we take the original starting dates and planned program durations of training as given. Thus, we first simulate the unemployment spells and then the employment spells, using those observations following simulated unemployment spells that end with a transition to employment. When we simulate earnings we use the estimated coefficients from our alternative model with earnings as the second outcome to simulate unemployment durations and daily earnings. We drop simulated observations for which the simulated unemployment spell with training ends before the actual start of training. In order to obtain the nontreatment outcomes we repeat the simulation procedure with treatment effects set to zero and using the same random draws. We compute standard errors by rerunning the simulation procedure 250 times using the parameter values from the bootstrap replications.

# I Selected Tests for Treatment Effect Heterogeneity

Table I1: Wald Tests of Heterogeneous Treatment Effects for Model 1 (Employment)

	$\chi^2$ -Statistic	df	<i>p</i> -Value
Unemployment to Employment			
Short-Term Training			
Female	7.027	3	0.071
Foreign	0.269	1	0.604
Age	13.524	2	0.001
Occupation	23.124	5	0.000
Long-Term Training			
Female	1.320	1	0.251
Foreign	13.566	5	0.019
Age	21.640	2	0.000
Education	30.006	3	0.000
Occupation	8.000	2	0.018
Employment to Unemployment			
Short-Term Training			
Female	0.827	1	0.363
Foreign	3.018	1	0.082
Occupation	0.991	1	0.320
Long-Term Training			
Female	4.150	2	0.126
Foreign	13.605	2	0.001
Education	5.130	2	0.077
Occupation	4.409	1	0.036

Notes: The hypothesis tests refer to joint significance tests on the corresponding estimated coefficients for model 1 in table F1. For instance, the test in the first row labeled ‘Female’ refers to the coefficients on the variables `f_stend31`, `f_stend361` and `pst_female` in model 1. The column ‘ $\chi^2$ -Statistic’ shows the test statistic, column ‘df’ the degrees of freedom, and ‘*p*-Value’ the probability value.



Table I2: Wald Tests of Heterogeneous Treatment Effects for Model 2 (Earnings)

	$\chi^2$ -Statistic	df	<i>p</i> -Value
Unemployment to Employment			
Short-Term Training			
Female	16.713	3	0.001
Foreign	4.782	1	0.029
Age	2.818	2	0.244
Occupation	19.027	5	0.002
Long-Term Training			
Female	8.562	1	0.003
Foreign	41.259	5	0.000
Age	11.108	2	0.004
Education	13.069	3	0.004
Occupation	0.094	2	0.954
Earnings			
Short-Term Training			
Female	15.095	1	0.000
Foreign	6.923	1	0.009
Occupation	0.896	1	0.344
Long-Term Training			
Female	11.522	1	0.001
Foreign	1.382	1	0.240
Occupation	22.472	2	0.000

Notes: The hypothesis tests refer to joint significance tests on the corresponding estimated coefficients for model 2 in table F1. For instance, the test in the first row labeled ‘Female’ corresponds to the coefficients on the variables `f_stend31`, `f_stend361` and `pst_female` in model 2. The column ‘ $\chi^2$ -Statistic’ shows the test statistic, column ‘df’ the degrees of freedom, and ‘*p*-Value’ the probability value.

## Additional References

- Bender, Stefan, Annette Haas, Christoph Klose (2000). "The IAB Employment Subsample 1975-1995." *Schmollers Jahrbuch*, 120 (4), 649-662.
- Bieback, Karl-Jürgen (1997). "Der Umbau der Arbeitsförderung." *Kritische Justiz*, 1/1997, 15-29.
- Blaschke, Dieter, and Hans-Eberhard Plath (2000). "Möglichkeiten und Grenzen des Erkenntnisgewinns durch Evaluation aktiver Arbeitsmarktpolitik." *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung (MittAB)*, 33 (3), 462-482.
- Bundesagentur für Arbeit (2001a, 2003a, 2005a). *Arbeitsmarkt 2000, 2002, 2004*. Nürnberg (various issues).
- Bundesagentur für Arbeit (2000, 2001b, 2002, 2003b, 2004, 2005b). *Daten zu den Eingliederungsbilanzen 1999, 2000, 2001, 2002, 2003, 2004*. Nürnberg (various issues).
- Bundesagentur für Arbeit (2005c). *Geschäftsbericht 2004*. Nürnberg.
- Eichhorst, Werner, Maria Grienberger-Zingerle, and Regina Konle-Seidl (2006). "Activation Policies in Germany: From Status Protection to Basic Income Support." IZA Discussion Paper No. 2514.
- Heckman, James J. and Christopher R. Taber (1994). "Econometric Mixture Models and more General Models for Unobservables in Duration Analysis." *Statistical Methods in Medical Research*, 3 (3), 279-299.
- Honoré, Bo (1990). "Identification Results for Duration Models with Multiple Spells or Time-Varying Covariates." Unpublished Manuscript, Northwestern University.
- Jacobebbinghaus, Peter and Stefan Seth (2007). "The German Integrated Employment Biographies Sample IEBS." *Journal of Applied Social Science Studies (Schmollers Jahrbuch)*, 127 (2), 335-342.
- Kühnlein, Gertrud and Birgit Klein (2003). "Bildungsgutscheine – mehr Eigenverantwortung, mehr Markt, mehr Effizienz? – Erfahrungen bei der Neuausrichtung der beruflichen Weiterbildung." Arbeitspapier Band 74, Hans-Böckler-Stiftung.

- Mosley, Hugh, Holger Schütz, and Günther Schmid, in collaboration with Kai-Uwe Müller, Stephan Manning, Sina Wilke and Sandra Gauss (2002). "Effizienzmobilisierung der Arbeitsverwaltung: Leistungsvergleich und Lernen von guten Praktiken (Benchmarking)." Wissenschaftszentrum Berlin für Sozialforschung, Discussion Paper FS I 02-209.
- OECD (2006). *OECD Employment Outlook*. Paris.
- Paulsen, Bent, Katrin Gutschow, Martin Isenmann, Angelika Puhlmann, Brigitte Gravalas and Frank Jander (2006). "Abschlussbericht zum Forschungsprojekt 3.5.101. Bildungsgutscheine in der öffentlich geförderten Weiterbildung: Erfahrungen und Auswirkungen." Bundesinstitut für Berufsbildung (BIBB).
- Platzmann, Gisela (2002). "Der Einfluss der Arbeitslosenversicherung auf die Arbeitslosigkeit in Deutschland." *Beiträge zur Arbeitsmarkt- und Berufsforschung (BeitrAB)*, 255.
- Schneider, Hilmar, Karl Brenke, Lutz Kaiser, Jakob Steinwedel, Birgit Jesske, and Arne Uhendorff (2006). "Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission – Modul 1b: Förderung beruflicher Weiterbildung und Transferleistungen." IZA Research Report No. 7, Bonn.
- Schweiger, Maximilian D. (2002). "Rechtliche Einordnung der durch das Job-AQTIV-Gesetz in das Arbeitsförderungsrecht eingefügten Eingliederungsvereinbarung (§35 Abs. 4 SGB III n.F.)." *Neue Zeitschrift für Sozialrecht*, 8/2002, 410-415.
- Seeleib-Kaiser, Martin and Timo Fleckenstein (2007). "Discourse, Learning and Welfare State Change: The Case of German Labour Market Reforms." *Social Policy and Administration*, 41 (5), 427-448.